

# ALGORITHMICALLY COMPLEX RESIDUALLY FINITE GROUPS

O. KHARLAMPOVICH<sup>1</sup>, A. MYASNIKOV<sup>2</sup>, AND M. SAPIR<sup>3</sup>

**ABSTRACT.** We construct the first examples of algorithmically complex finitely presented residually finite groups, the first examples of finitely presented residually finite groups with arbitrarily high (recursive) Dehn functions, and arbitrary large depth functions. The groups are solvable of class 3. We also prove that the universal theory of finite solvable of class 3 groups is undecidable.

## CONTENTS

1. Introduction	2
1.1. The problem and previous approaches for a solution	2
1.2. The “yes” and “no” parts of the McKinsey algorithm	4
1.3. Quantification of the “yes” part: the Dehn function	5
1.4. Quantification of the “no” part: the depth function	6
1.5. Distortion of pro-finitely closed subgroups of finitely presented groups	7
1.6. Methods of proof	8
1.7. Structure of the paper	8
2. Turing machines and Minsky machines	9
2.1. Turing machines	9
2.2. Universally halting Turing machines	10
2.3. Minsky machines	12
3. Simulation of Minsky machines by semigroups	13
3.1. The construction	13
3.2. Residually finite finitely presented semigroups	15
3.3. Residually finite semigroups with large depth function	16
4. Simulation of Minsky machines in solvable groups	18
4.1. The construction	19
4.2. A finitely presented solvable group with undecidable word problem	25
4.3. Residually finite finitely presented groups	26
4.4. Residually finite finitely presented group with large depth function	26
5. Distortion of subgroups closed in the pro-finite topology	27
6. Universal theories of sets of finite solvable groups	28
References	29

---

<sup>1</sup> Partially supported by NSF grant DMS-0700811, <sup>2</sup> Partially supported by NSF grants DMS-0700811 and DMS-0914773. <sup>3</sup> Partially supported by NSF grant DMS-0700811 and by BSF (U.S.A.-Israel) grant 2010295.

## 1. INTRODUCTION

**1.1. The problem and previous approaches for a solution.** It is well known that finitely presented residually finite groups are much simpler algorithmically than arbitrary finitely presented groups. For example, the word problem in every such group is decidable. Moreover the most “common” residually finite groups, the linear groups over fields, are algorithmically “tame”: the word problem in any linear group is decidable in polynomial time and even log-space [35]. Dehn functions are important witnesses of the complexity of the word problem. Although it is well known that groups with solvable word problem can have very large Dehn functions [40, 58], no examples of finitely presented residually finite groups with superexponential Dehn function are known. Thus one of the main open problems in this area is how large could the Dehn function of a residually finite finitely presented group be. The question was known since early 90s. It was open for so long because all known methods to construct algorithmically hard groups produced either non-residually finite groups or groups where the question about their residual finiteness is very difficult. Not much is known even for linear groups (Steve Gersten asked [19, 20] if there exists a uniform upper bound for Dehn functions of linear groups). Let us briefly discuss the previous attempts to solve the problem and the reasons why these methods did not work.

**1.1.1. Method 1. Known groups with large Dehn functions.** One could hope that some of the known finitely presented groups with very large Dehn function may turn out to be residually finite, which would shed some light on how to produce residually finite groups with even larger Dehn functions. Unfortunately, this is not the case, all these groups are non-residually finite. For example, the Dehn function of the one relator group  $G_{(1,2)} = \langle a, b \mid b^{-1}a^{-1}bab^{-1}ab = a^2 \rangle$  introduced by Baumslag in [2] is bigger than any iterated exponent (see Gersten [18]). Platonov [51] proved that it is equivalent to the function  $\exp^{(\log 2n)}(1)$ , where  $\exp^{(m)}(x)$  is the function defined by  $\exp^{(0)}(n) = n$  and  $\exp^{(k+1)}(n) = \exp(\exp^{(k)}(n))$ . However,  $G_{(1,2)}$  is not residually finite (and in fact has very few finite quotients) [2]. Furthermore, the word problem in  $G_{(1,2)}$  is in polynomial time [42].

**1.1.2. Method 2. Using subgroups with very large distortion.** Consider a finitely presented group  $G$  and a “badly” distorted finitely generated subgroup  $H$ . Let  $T = \langle G, t \mid t^{-1}ht = h \ (h \in H) \rangle$  be the HNN extension of  $G$  where the free letter  $t$  centralizes  $H$ . It was noticed by Bridson and Häfliger [11, Theorem 6.20.III] that the Dehn function of  $T$  is at least as large as the distortion function of  $H$  in  $G$ . The following result puts some limitations on this method of constructing complicated residually finite groups.

**Lemma 1.1.** *If the group  $T$  is residually finite then  $H$  is closed in the pro-finite topology of  $G$ .*

*Proof.* In the notation above, suppose  $T$  is residually finite and  $u \in G \setminus H$ . Then  $w = [u, t] \neq 1$  in  $T$  (by the standard properties of HNN extensions). Hence there exists a homomorphism  $\phi$  from  $T$  onto a finite group  $T_w$  such that  $\phi(w) \neq 1$ . But this implies  $\phi(u) \notin \phi(H)$  (since every element of  $\phi(H)$  commutes with  $\phi(t)$ ). Hence there exists a normal subgroup of finite index  $N < G$  such that  $u$  does not belong to  $NH$ . In other words  $H$  must be closed in the pro-finite topology of  $G$ .  $\square$

Consider the following typical examples of residually finite groups  $G$  with highly distorted subgroups  $H$ . The first one is Wise’s version [64] of Rips’ construction [53] which for every

finitely presented group  $Q$  gives a finitely presented residually finite small cancellation group  $G$  with a short exact sequence

$$1 \rightarrow N \rightarrow G \rightarrow Q \rightarrow 1$$

where  $N$  finitely generated. It is easy to see that the distortion function of  $N$  in  $G$  is at least as bad as the Dehn function in  $Q$ , so choosing  $Q$  properly one can get a finitely presented residually finite group  $G$  with a highly distorted subgroup  $N$ . Now, the subgroup  $N$  is normal in the HNN extension  $T$ . So it is closed in the pro-finite topology of  $T$  only if  $Q = G/N$  is residually finite. By Lemma 1.1,  $T$  can be residually finite only if  $Q$  is residually finite. In other words to construct a complicated finitely presented residually finite group  $T$  one has to have the initial group  $Q$  complicated, finitely presented and residually finite as well.

The second example is the standard Mikhailova construction. In this case highly distorted subgroups of the direct product of two free groups  $F_2 \times F_2$  can be obtained as equalizers of two homomorphisms  $\phi_1: E_1 \rightarrow M$  and  $\phi_2: E_2 \rightarrow M$  where  $E_1, E_2$  are finitely generated subgroups of  $F_2$  and  $M$  is finitely presented (see Sections 1.4, 5 below). But by Remark 5.4 below the equalizer is closed in the pro-finite topology only if  $M$  is residually finite. Thus as in the previous example, in order to construct an algorithmically complex residually finite finitely presented group using Mikhailova's construction and the HNN extensions as above, one needs to have already a finitely presented residually finite algorithmically complex group  $M$ .

The third example is Cohen's construction of highly distorted subgroups employing the modular machines [12]. One can also prove that in that construction the subgroup will be pro-finitely closed only if the modular machine is very easy.

One can also try to use the hydra groups [14, 10] to construct HNN extensions as above with Dehn functions bigger than any prescribed Ackermann function. The question of whether these groups are residually finite is open (most probably the answer is negative because the distorted subgroup should not be closed in the pro-finite topology, but this needs a proof).

**1.1.3. Method 3. Boone-Novikov constructions.** One of the standard ways to produce algorithmically complicated groups is by simulating Turing machines using free constructions (HNN extensions and amalgamated products) which goes back to the seminal papers by Boone and Novikov (see, for example, [54]). There are currently many versions of that construction (for a recent survey see [57]). But in fact, it can be shown that for each known version of the proof of Boone-Novikov theorem using free constructions, even for easy Turing machines the corresponding group is non-residually finite. This is, for instance, the main idea of the example in [31]. Here is an even easier example. Let  $G = G(M)$  be a group constructed by any of these constructions. Then for every input word  $u$  of the Turing machine  $M$  there exists a word  $w = w(u)$  obtained by inserting some copies of  $u$  in  $w(\emptyset)$ , so that  $u$  is accepted by  $M$  if and only if  $w(u) = 1$  in  $G$ . Now consider  $M$  that accepts a word  $a^n$  if and only if  $n \neq 0$  (that machine is actually one of the basic building blocks in [58]). Then  $w(a^n) = 1$  in  $G$  if and only if  $n \neq 0$ . Suppose that there exists a homomorphism  $\phi$  onto a finite group  $H$  that separates  $w(\emptyset) = w(a^0)$  from 1. Then  $\phi(a)$  has finite order, say,  $s$ , in  $H$ . Therefore  $\phi(w(\emptyset)) = \phi(w(a^0)) = \phi(w(a^s)) = \phi(1) = 1$ , a contradiction. Hence  $G(M)$  is not residually finite.

**1.1.4. Method 4. Residually finite groups obtained by free constructions.** In general, the question about residual finiteness of free constructions is very difficult. Currently there are only two large classes of groups where the question was settled: these are ascending HNN extensions of free groups [7] and certain groups acting “nicely” on CAT(0)-cubical complexes including small cancelation groups (see the recent work of Wise, for example, [25] and references therein). All these groups have easy word problem and uniformly bounded Dehn functions. The reason for the lack of more examples is that groups obtained by free constructions from “nice” groups contain a lot of extra elements and it is not at all clear how to separate these elements from 1 by homomorphisms onto finite groups. In the two cases when it could be done, it was possible to reformulate the problem in the language of algebraic geometry and geometric topology, respectively.

**1.1.5. Method 5. Other groups with complicated word problem.** There are several other constructions of groups with complicated word problem but each of these also almost always produce non-residually finite groups. For example, the group in [38] is based on the R. Thompson group  $V$  which is infinite and simple (hence not residually finite).

**1.1.6. The main results of the paper.** In this paper, we construct finitely presented residually finite groups with arbitrary complex word problem, and also easy word problem but arbitrary high (of course recursive) Dehn function, and arbitrary high depth function. We also give applications of these results to the questions about solvability of the universal theory of finite solvable groups.

Our constructions are based on simulating Minsky machines in groups. Surprisingly, the algebraic structure of the group not only depends on the construction itself, but also heavily depends on the computational properties of the machines we simulate: for the group to be residually finite the Minsky machines should be sym-universally halting, that is their transition graphs are vertex-disjoint unions of finite trees.

**1.1.7. What next?** We expect the approach used in this paper to be useful in solving other problems that are still open. For example, the residually finite version of the Higman embedding theorem would be very desirable. It is known [39, 15, 21, 33] that a finitely generated recursively presented residually finite group may have undecidable word problem, and hence cannot be embedded into a finitely presented residually finite group. But it is not known whether every unsolvability of the word problem is the only obstacle for such an embedding. Thus it would be very interesting to find out whether every finitely generated residually finite group with solvable word problem embeds into a finitely presented residually finite group. Note that usually a version of Boone-Novikov theorem precedes a version of Higman theorem, hence we can consider this paper as a step toward the residually finite version of Higman’s theorem.

**1.2. The “yes” and “no” parts of the McKinsey algorithm.** One of the initial motivations for studying residually finite groups, semigroups and other algebraic structures was McKinsey’s algorithm solving the word problem in finitely presented residually finite algebraic structures. Even though the algorithm is well known and classical, surprisingly little is known about its complexity. In this paper, we fill this gap.

Let  $G = \langle X; R \rangle$  be a residually finite finitely presented algebraic structure of finite type (signature)  $T$  (say, groups, semigroups, rings, etc.) Let us recall McKinsey’s algorithm solving the word problem in  $G$  (see [41], [37]). The word problem is divided into two parts.

Let  $F(X)$  be the free algebraic structure of type  $T$  freely generated by  $X$ . Then we define the “yes” and “no” parts of the word problem in  $G$  as follows:

$$\text{WP}_{\text{yes}} = \{(w, w') \in F(X) \mid w =_G w'\} \text{ and } \text{WP}_{\text{no}} = \{(w, w') \in F(X) \mid w \neq_G w'\}.$$

To solve the word problem in  $G$  one runs in parallel two separate algorithms  $\mathcal{A}_{\text{yes}}$  and  $\mathcal{A}_{\text{no}}$ , such that starting with a given pair of elements  $w, w' \in F(X)$   $\mathcal{A}_{\text{yes}}$  stops if and only if  $(w, w') \in \text{WP}_{\text{yes}}$  and  $\mathcal{A}_{\text{no}}$  stops if and only if  $(w, w') \in \text{WP}_{\text{no}}$ .

**The algorithm  $\mathcal{A}_{\text{yes}}$**  enumerates one by one all consequences of the defining relations  $R$  and waits until  $w = w'$  appears in the list.

**The algorithm  $\mathcal{A}_{\text{no}}$**  enumerates all homomorphisms  $\phi_1, \phi_2, \dots$ , of  $G$  into finite algebraic structures of type  $T$  and waits until  $\phi_i(w) \neq \phi_i(w')$ .

Let now  $G$  be a finitely presented residually finite group. Although it seems like in general  $\mathcal{A}_{\text{yes}}$  and  $\mathcal{A}_{\text{no}}$  are very slow, there were no examples of groups  $G$  for which these algorithms were actually very slow. More precisely, there were no known examples of finitely presented residually finite groups with very hard “yes” or “no” part of the word problem. Indeed, the most “common” residually finite groups are linear groups, say, over fields [37]. In that case it is well known that the “yes” part can be solved in deterministic polynomial time [35, 61]. The “no” part can be solved by considering factor groups corresponding to ideals of finite index of some polynomial rings, hence also can be shown to be solvable in deterministic polynomial time. In fact the same can be said about most finitely presented groups (where “most” means “with overwhelming probability” in one of several probabilistic models): recent results of Agol [1] and Ollivier and Wise [47] together with the older result of Olshanskii [48] imply that most finitely presented groups are linear (even over  $\mathbb{Z}$ ).

One of our main results is the following theorem (an immediate corollary of Theorem 4.19 below):

**Theorem.** *Let  $f(n)$  be a recursive function. Then there exists a residually finite finitely presented solvable group  $G$  such that for any finite presentation  $\langle X; R \rangle$  of  $G$  the time complexity of both “yes” and “no” parts of the word problem are at least as high as  $f(n)$ .*

We also show that both algorithms  $\mathcal{A}_{\text{yes}}$  and  $\mathcal{A}_{\text{no}}$  can be very slow even when both “yes” and “no” parts of the word problem are easy.

*Remark 1.2.* Note that if we replace “finitely presented” assumption by “recursively presented”, then residually finite groups are known to be very complicated. As we have mentioned before, recursively presented finitely generated residually finite groups may have undecidable word problem [39, 15, 21]). Even more, recently the second author and B. Khoussainov constructed residually finite Dehn monsters, i.e., infinite groups which are recursively presented, residually finite and algorithmically finite [33]. These are groups where the word problem is not only undecidable, but one cannot algorithmically enumerate an infinite set of pair-wise distinct elements of the group.

Note also that although our groups are not linear they are (elementary Abelian)-by-linear since they are solvable of class 3 with the second derived subgroup elementary Abelian.

**1.3. Quantification of the “yes” part: the Dehn function.** It was noticed by Madlener and Otto [40] that the Dehn function of a group measures the complexity of the word problem. They also constructed finitely presented groups with arbitrary large Dehn functions. For residually finite groups, the situation is different. Nilpotent groups are examples of

residually finite groups with arbitrary high polynomial Dehn function [4]. The Baumslag-Solitar groups  $\langle x, y \mid x^y = x^k \rangle$ ,  $k \geq 2$ , are examples of residually finite (even linear) groups with exponential Dehn function. No examples of residually finite groups with bigger Dehn functions were known. This gap is filled by the following

**Theorem 4.18.** *For every recursive function  $f$ , there is a residually finite finitely presented solvable of class 3 group  $G$  with Dehn function greater than  $f$ . In addition, one can assume that the word problem in  $G$  is at least as hard as any given recursive function or as easy as polynomial time.*

As a corollary of Theorem 4.18 we mention the following exotic examples of groups.

**Corollary.** *For every recursive function  $f$ , there is a residually finite finitely presented solvable of class 3 group  $G$  with Dehn function greater than  $f$  and the word problem decidable in polynomial time.*

**1.4. Quantification of the “no” part: the depth function.** The function quantifying the algorithm  $\mathcal{A}_{\text{yes}}$  is the *depth function* introduced by Bou-Rabee [9]. Recall that if  $G = \langle X \rangle$  is a finitely generated group or semigroup, the depth function  $\rho_G(n)$  is the smallest function such that every two words  $w \neq_G w'$  of length at most  $n$  are separated by a homomorphism to a group (semigroup)  $H$  with  $|H| \leq \rho_G(n)$ . That function does not depend on the choice of finite generating set  $X$  (up to the natural equivalence).

It is easy to see that for every finitely generated linear group or semigroup  $G$ ,  $\rho_G$  is at most polynomial. Since finitely generated metabelian groups are subgroups of direct products of linear groups [62] the depth function of every finitely generated metabelian group is at most polynomial. By the recent result of Agol [1] based on the earlier results of Wise [63], every small cancellation group is a subgroup of a Right Angled Artin group, hence linear and has polynomial depth function. In fact one can have much smaller bounds for many linear groups. For example, for the free group  $F_2$ ,  $\rho_{F_2}(n)$  is at most  $n^{\frac{2}{3}}$  by a result of Kassabov and Matucci [26]. There are some finitely presented groups for which the depth function is unknown and very interesting. For example the ascending HNN extensions of free groups are known to be residually finite and even virtually residually nilpotent (proved by A. Borisov and the third author [7, 8]) but the only upper bound one can deduce from the proof is exponential. Although many of these groups have small cancellation presentations and so covered by the results from [1], there are some groups of this kind for which the depth function is not known. One of these groups is  $\langle x, y, t \mid txt^{-1} = xy, tyt^{-1} = yx \rangle$ . The fact that it is hyperbolic follows from Bestvina-Feign combination theorem [6] and was proved by Minasyan (unpublished). If the depth function of that group is not polynomial, that group would not be linear, disproving a conjecture by Wise (he conjectured that all hyperbolic ascending HNN extensions of free groups are linear and, moreover, subgroups of Right Angled Artin groups).

For finitely generated infinitely presented groups (even amenable ones) the situation is much more clear now. Using the method of Kassabov and Nikolov [27] and the result of Nikolov and Segal [46] one can construct a finitely generated residually finite group with arbitrary large recursive depth function.

In this paper, we show that a similar result holds for finitely presented solvable of class 3 groups.

**Theorem 4.19.** *For every recursive function  $f$ , there is a residually finite finitely presented solvable of class 3 group  $G$  with depth function greater than  $f$ . In addition, one can assume*



that the word problem in  $G$  is at least as hard as the membership problem in a given recursive set of natural numbers  $Z$  or as easy as polynomial time.

As a corollary of Theorem 4.20 we mention the following exotic examples of groups.

**Corollary.** *For every recursive function  $f$ , there is a residually finite finitely presented solvable of class 3 group  $G$  with depth function greater than  $f$  and the word problem decidable in polynomial time.*

**1.5. Distortion of pro-finitely closed subgroups of finitely presented groups.** Let  $G$  be a group generated by a finite set  $X$ ,  $H \leq G$  be a subgroup generated by a finite set  $Y$ . Recall that the distortion function  $f_{H,G}(n)$  is defined as the minimal number  $f$  such that every element of  $H$  represented as a word  $w$  of length  $\leq n$  in the alphabet  $X$  can be represented as a word of length  $\leq f$  in the alphabet  $Y$  [16]. It is clear [16] that the distortion function  $f_{G,H}$  is recursive if and only if the membership problem in  $H$  is decidable.

Recall that  $H$  is *closed in the pro-finite topology* of  $G$  if  $H$  is the intersection of some subgroups of  $G$  of finite index. If  $G$  is finitely presented and  $H$  is closed in the pro-finite topology of  $G$ , then there exists a McKinsey-type algorithm  $A(G, H)$  solving the membership problem for  $H$  (and thus the  $f_{G,H}$  is recursive). For every word  $w$  in the alphabet  $X$ , the “yes” part  $A_{\text{yes}}(G, H)$  of the algorithm lists all words in  $Y$ , rewrites them as words in  $X$ , and then applies relations of  $G$  to check whether one of these words is equal to  $w$ . The “no” part  $A_{\text{no}}(G, H)$  of the algorithm lists all homomorphisms  $\phi$  of  $G$  into finite groups and checks whether  $\phi(w) \notin \phi(H)$ . As in Section 1.2, one can ask what is the complexity of the “yes” and “no” parts of that algorithm, in particular, and of the membership problem for  $H$  in general.

One can also quantify the complexity of the two parts  $A_{\text{yes}}(G, H)$  and  $A_{\text{no}}(G, H)$ . The “yes” part is quantified by the distortion function  $f_{G,H}(n)$  and the “no” part is quantified by the *relative depth function*  $\rho_{G,H}(n)$  which is defined as the minimal number  $r$  such that for every word  $w$  of length  $\leq n$  in  $X$  which does not represent an element of  $H$  there exists a homomorphism  $\phi$  from  $G$  to a finite group of order  $\leq r$  such that  $\phi(w) \notin \phi(H)$ .

As for the word problem in residually finite finitely presented groups (discussed above), there were no examples of finitely generated subgroups of finitely presented groups that are closed in the pro-finite topology but have “arbitrary bad” distortion or “arbitrary bad” relative depth function.

The Mikhailova’s construction shows that finitely generated subgroups of the residually finite group  $F_2 \times F_2$  (here  $F_2$  is a free group of rank 2) could be as distorted as one pleases. In fact the set of possible distortion functions of subgroups of  $F_2 \times F_2$  coincides, up to a natural equivalence, with the set of Dehn functions of finitely presented groups [49]. By a result of Baumslag and Roseblade [5] subgroups of  $F_2 \times F_2$  are *equalizers* of pairs of homomorphisms  $\phi: F_k \rightarrow G, \psi: F_n \rightarrow G$  (where  $F_k, F_n$  are subgroups of  $F_2$ ), i.e. the subgroups of the form  $\{(x, y) \in F_k \times F_n \mid \phi(x) = \psi(y)\}$ . The equalizer subgroup is finitely generated if and only if  $G$  is finitely presented. It is easy to prove (see Lemma 5.2 below) that if  $G$  is residually finite, then the equalizer is closed in the pro-finite topology of  $F_2 \times F_2$ . Thus we can use the examples of residually finite finitely presented groups with complicated word problem and complicated depth function to prove the following

**Theorem 5.5** *For every recursive function  $f(n)$  there exists a finitely generated subgroup  $H \leq F_2 \times F_2$  that is closed in the pro-finite topology of  $F_2 \times F_2$  and whose distortion function*

$f_{F_2 \times F_2, H}$ , the relative depth function, and the time complexities of both “yes” and “no” parts of the membership problem are at least  $f(n)$ .

There is an analogous (though a bit weaker) result, for subgroups of a direct product  $S_3(X) \times S_3(X)$ , where  $S_3(X)$  is a free solvable group of class 3 with free generating set  $X$ .

**Theorem 5.6** *For any recursive function  $f(n)$  there is a finite set  $X$  and a finitely generated subgroup  $H \in S_3(X) \times S_3(X)$  such that  $E$  is closed in the pro-finite topology on  $S_3(X) \times S_3(X)$  and the distortion function, the relative depth function, and the time complexities of both “yes” and “no” parts of the membership problem in  $H$  are at least  $f(n)$ .*

**1.6. Methods of proof.** As we have shown above (see Section 1.1.3) most versions of the Boone-Novikov construction ([54, 38, 12, 58]) do not produce residually finite groups. Instead, we simulate Minsky machines in solvable groups of class 3. A similar construction is due to the first author [28]. We use a version from the unpublished thesis [29]. As in [28, 29], our group is a split extension of an elementary Abelian group of prime exponent by a metabelian group. Since every metabelian group has easy word problem and is residually finite, we can concentrate only on the elementary Abelian subgroup which is spanned, basically, by the configurations of the Minsky machine encoded in a certain way. That encoding is a very important feature of the construction. It helps us avoid the problem from Section 1.1.3 because the words  $w(u)$  corresponding to the input configurations  $u$  of the Minsky machines are not obtained by inserting copies of  $u$  in  $w(\emptyset)$ .

Of course that construction also often leads to non-residually finite groups. But it turned out that the difficulty can be overcome by modifying the Minsky machine first. In this paper, we use the fact that every Turing machine and every Minsky machine with decidable halting problem is equivalent to a *universally halting* and even *sym-universally halting* machine (these machines can be characterized as the machines whose transition graphs are vertex disjoint unions of finite trees).

**1.7. Structure of the paper.** The paper is organized as follows. Section 2 contains preliminary results about Turing and Minsky machines that are needed further. We show that one can modify any Turing or Minsky machine that recognizes a recursive set into a machine that halts on every configuration. In fact we can even assume that the symmetrized machine always halts (we call such machines sym-universally halting).

In Section 3, we simulate sym-universally halting Minsky machines in residually finite finitely presented semigroups and prove the analogs of the above theorems for semigroups.

In Section 4 we simulate Minsky machines in solvable groups and construct complicated residually finite finitely presented groups.

Sections 5 and 6 contain applications of the main theorems. In Section 5 we prove, in particular, Theorem 5.5. In Section 6, we strengthen the well-known result of Slobodskoi about undecidability of the universal theory of finite groups. We show, in particular, that the universal theory of any set of finite groups that contains all finite solvable groups of class 3 is undecidable.

**Acknowledgement.** The authors are grateful to Jean-Camille Birget and Friedrich Otto for pointing to the references [13], to Ben Steinberg for pointing to the reference [35], to Rostislav Grigorchuk for pointing to the references [15, 21] and to Tim Riley for pointing to the references [19, 20]. We are also grateful to Markus Lohrey and Ralph Strebel for their comments.



## 2. TURING MACHINES AND MINSKY MACHINES

**2.1. Turing machines.** Let us give a definition of a Turing machine. A Turing machine  $M$  with  $K$  tapes consists of hardware (the tape alphabet  $A = \sqcup_{i=1}^k A_i$ , and the state alphabet  $Q = \sqcup_{i=1}^K Q_i$ <sup>1</sup>) and program  $P$  (the list of commands, defined below). A *configuration* of a Turing machine  $M$  is a word

$$\alpha_1 u_1 q_1 v_1 \omega_1 \quad \alpha_2 u_2 q_2 v_2 \omega_2 \quad \dots \quad \alpha_K u_K q_K v_K \omega_K$$

(we included spaces to make the word more readable) where  $u_i, v_i$  are words in  $A_i$ ,  $q_i \in Q_i$  and  $\alpha_i, \omega_i$  are special symbols (not from  $A \cup Q$ ).

A command simultaneously replaces subwords  $a_i q_i b_i$  by words  $a'_i q'_i b'_i$  where  $a_i, a'_i$  are either letters from  $A_i \cup \{\alpha_i\}$  or empty,  $b_i, b'_i$  are either letters from  $A_i \cup \{\omega_i\}$  or empty. A command cannot insert or erase  $\alpha_i$  or  $\omega_i$ , so if, say,  $a_i = \alpha_i$ , then  $a'_i = \alpha_i$ . Note that with every command  $\theta$  one can consider the *inverse* command  $\theta^{-1}$  which undoes what  $\theta$  does.

A *computation* of  $M$  is a sequence of configurations and commands from  $P$ :

$$w_1 \xrightarrow{\theta_1} w_2 \xrightarrow{\theta_2} \dots \xrightarrow{\theta_l} w_{l+1}.$$

Here  $l$  is called the *length* of the computation. We choose *stop states*  $q_i^0$  in each  $Q_i$ , then we can call a configuration  $w$  *accepted* if there exists a computation starting with  $w$  and ending with a configuration where all state symbols are  $q_i^0$  and all tapes are empty. Also we choose *start states*  $q_i^1$  in each  $Q_i$ . Then an *input* configuration corresponding to a word  $u$  over  $A_1$  is a configuration  $\text{inp}(u)$  of the form

$$\alpha_1 u q_1^1 \omega_1 \quad \alpha_2 q_2^1 \omega_2 \quad \dots \quad \alpha_K q_K^1 \omega_K.$$

We say that a word  $u$  over  $A_1$  is accepted by  $M$  if the configuration  $\text{inp}(u)$  is accepted. The set of all words accepted by  $M$  is called the *language accepted by  $M$* .

The *time function*  $T_M(n)$  of  $M$  is the minimal function such that every accepted word of length  $\leq n$  has an accepting computation of length  $\leq T_M(n)$ . The *space function*  $S_M(n)$  of  $M$  is the minimal function such that every accepted word of length  $\leq n$  has an accepting computation where every configuration has length  $\leq S_M(n)$ .

A Turing machine  $M$  is called *deterministic* if for every configuration, there exists at most one command from the program  $P$  that applies to this configuration.

In this paper, we shall consider several types of machines. A machine  $M$  in general has an alphabet and a set of words in that alphabet called *configurations*. It also has a finite set of commands. Each command is a partial injective transformation of the set of configurations. A computation is a sequence

$$w_1 \xrightarrow{\theta_1} w_2 \xrightarrow{\theta_2} \dots \xrightarrow{\theta_l} w_{l+1}.$$

where  $w_j$  are configurations,  $\theta_1, \dots, \theta_n$  are commands and  $\theta_i(w_i) = w_{i+1}$  for every  $i = 1, \dots, n$ . A machine is called *deterministic* if the domains of its commands are disjoint. A machine usually has a distinguished *stop* configuration, and a set  $I = I(M)$  of *input* configurations. A configuration is called *accepted* by  $M$  if there exists a computation connecting that configuration with the stop configuration. The machine  $\text{Sym}(M)$  is defined in the natural way (add the inverses of all commands of  $M$ ). Two configurations  $w, w'$  are called *equivalent*, written  $w \equiv_M w'$ , if there exists a computation of  $\text{Sym}(M)$  connecting these configurations. Clearly,  $\equiv_M$  is an equivalence relation.

---

<sup>1</sup> $\sqcup$  denotes disjoint union

The following general lemma is easy but useful.

**Lemma 2.1.** *Suppose that  $M$  is deterministic. Then two configurations  $w, w'$  of  $M$  are equivalent if and only if there exists two computations of  $M$  connecting  $w, w'$  with the same configuration  $w''$  of  $M$ .*

*Proof.* Indeed, since  $M$  is deterministic, in any computation of  $\text{Sym}(M)$  where no command is followed by its inverse inverses of command of  $M$  cannot be followed by commands of  $M$ . Thus the computation is a concatenation of two (possibly empty) parts: the first part uses only commands of  $M$ , the second part uses only inverses of commands of  $M$ .  $\square$

We say that a set  $X$  of natural numbers is *enumerated* by a machine  $M$  if there exists a recursive encoding  $\mu$  of natural numbers by input configurations of  $M$  such that a number  $u$  belongs to  $X$  if and only if  $\mu(u)$  is accepted by  $M$ . The set  $X$  is *recognized* by  $M$  if  $M$  enumerates  $X$  and for every input configuration every computation starting with that configuration eventually halts (arrives to a configuration to which no command of  $M$  is applicable).

We say that machine  $M'$  *polynomially reduces* to a machine  $M$  if there exists an polynomial time algorithm  $A$  checking equivalence of configurations of  $M'$  which uses an oracle checking equivalence of configurations of  $M$  such that

- Any computation of  $A$  verifying equivalence of configurations  $c, c'$  of  $M'$  involves at most polynomial (in terms of  $|c| + |c'|$ ) number of uses of the oracle,
- and every time the sizes of the configurations of  $M$  whose equivalence the oracle should check are polynomially bounded in terms of  $|c| + |c'|$ .

We say that  $M$  and  $M'$  are *polynomially equivalent* if there are polynomial reductions of  $M$  to  $M'$  and vice versa.

**2.2. Universally halting Turing machines.** A (not necessarily deterministic) machine  $M$  is called *universally halting* if for every configuration  $w$  of  $M$  there exist only finitely many computations of  $M$  starting with  $w$  without repeated configurations.

We call a deterministic machine  $M$  *sym-universally halting* if  $\text{Sym}(M)$  halts if it starts with any non-accepted configuration.

**Theorem 2.2** (See, for example, [13]). *For every recursive set  $X$  of natural numbers, that is accepted by a deterministic Turing machine  $M$  there exists a universally halting deterministic Turing machine  $M'$  with one tape accepting  $X$  and polynomially equivalent to  $M$ .*

**Lemma 2.3.** *Let  $M$  be a deterministic sym-universally halting Turing machine. Then there exists a one-tape deterministic sym-universally halting Turing machine  $M'$  recognizing the same language as  $M$ . The machine  $M'$  is polynomially equivalent to  $M$ .*

*Proof.* The proof is by inspection of the proof from [17].  $\square$

**Theorem 2.4.** *For every recursive set of natural numbers  $X$  there exists a sym-universally halting Turing machine  $M''$  with one tape that recognizes  $X$ . The machine  $M''$  satisfies the following conditions.*

- (a) *For every configuration  $c$  of  $M''$  either  $c$  is equivalent to an input configuration or every computation of  $\text{Sym}(M'')$  starting with  $c$  has length at most  $O(|c|)$ .*
- (b) *If  $c, c'$  are two distinct input configurations of  $M''$  such that  $c \equiv_{M''} c'$ . Then either  $c = c'$  or both  $c, c'$  are accepted by  $M''$ .*

(c) If  $M$  is any Turing machine recognizing  $X$  then we can assume that  $M''$  polynomially reduces to  $M$ .

*Proof.* Let  $M$  be a deterministic universally halting Turing machine with  $K$  tapes recognizing  $L$ . Consider a new Turing machine  $M'$  constructed as follows.  $M'$  has one more tape than  $M$ , called the *history* tape. Its alphabet  $A'$  is in one-to-one correspondence with the set of commands  $P$  of  $M$ :  $A' = \{[\theta], \theta \in P\}$ . Its state alphabet consists of two letters  $q_{K+1}^0$  and  $q_{K+1}^1$ . With every command  $\theta$  of  $M$  we associate the command  $\theta'$  of  $M'$ . It does what  $\theta$  would do on the first  $k$  tapes of  $M'$  and inserts  $[\theta]$  on the history tape of  $M$ . After the first  $K$  tapes of  $M'$  form an accept configuration, the machine erases the history tape and stops (turns  $q_{K+1}^1$  into the stop state  $q_{K+1}^0$ ). Let  $P'$  be the program of  $M'$ . Now modify  $M'$  further to obtain a new Turing machine  $M''$ . The program  $P''$  of  $M''$  contains a copy  $\tilde{P}$  (the set of the *main commands*) of  $P'$  and some new, auxiliary, commands. After each main command of  $\tilde{P}$ ,  $M''$  executes the history written on the history tape backward, without erasing the history tape: it just scans the tape from left to right, reading the symbols written there one by one and executing on the first  $K$  tapes the inverses of the commands written on the history tape. The commands that do that will be called *auxiliary*. If at the end of the scanning the history tape, it reaches an input configuration,  $M''$  executes on the first  $K$  tapes the history written on the history tape in the natural order (scanning the history tape from right to left). After that  $M''$  is ready to execute the next main command. We do not give precise definition of the program of  $M''$  because it is obvious on the one hand and long on the other hand. Clearly, the state alphabet of  $M''$  must be bigger than the state alphabet of  $M'$ . The machine  $M''$  is deterministic, universally halting, and recognizes the same language  $L$ .

Let us prove properties (a) and (b) of the theorem. Since  $M''$  is deterministic, every reduced (i.e. without mutually inverse consecutive commands) computation  $\Theta$  of  $\text{Sym}(M'')$  is of the form  $\Theta_1\Theta_2^{-1}$  for some computations  $\Theta_1, \Theta_2$  of  $M''$  (because a command of  $(M'')^{-1}$  cannot be followed by its inverse).

Let us show that  $\text{Sym}(M'')$  halts when it starts with any non-accepted configuration (and then apply Lemma 2.3). Let  $w$  be a configuration of  $M''$  that is not accepted by  $M''$ . Since  $M''$  is deterministic, every computation of  $\text{Sym}(M'')$  starting at  $w$  is a concatenation of a computation of  $M''$  followed by a computation of  $(M'')^{-1}$  (i.e. the machine  $M''$  where every command is replaced by its inverse). Since  $M$  is universally halting, there are only finitely many computations of  $M''$  starting with  $w$ . Thus we only need to show that there are finitely many computations of  $(M'')^{-1}$  starting with  $w$ , or, equivalently, that there are only finitely many computations of  $M''$  ending with  $w$ . Suppose that there are infinitely many computations of  $M''$  ending with  $w$ . Then, by definition of  $M''$  there must exist infinitely many input configuration  $\text{inp}(u)$  of  $M''$  for which there exists a computation of  $M''$  starting with  $\text{inp}(u)$  and ending at  $w$ . But that is impossible because such  $\text{inp}(u)$  is unique and is obtained by applying the inverse of the history written on the history tape of  $M''$  to  $w$ .

(c) The fact that  $M''$  polynomially reduces to  $M$  is proved as follows. Consider two configurations  $w, w'$  of  $M''$ . If  $w$  is not equivalent to an input configuration, then by (b) we need to check only whether  $w'$  is one of  $O(|w|)$  words that belong to the longest computation of  $\text{Sym}(M'')$  containing  $w$ . That can be done in polynomial time without using the oracle checking equivalence of configurations of  $M$ . Suppose that both  $w$  and  $w'$  are equivalent to input configurations  $u, u'$  of  $M''$ . Then we can find  $u, u'$  in polynomial time and their lengths at at most  $O(|w| + |w'|)$ . If  $u \neq u'$  and either  $u$  or  $u'$  is not accepted, then by (b)

$w$  is not equivalent to  $w$ . If  $u = u'$ , then  $w$  is equivalent to  $u$ . Thus  $w$  is equivalent to  $w'$  if and only if  $u$  and  $v$  are accepted. To check that  $u$  is accepted, we need to remove letters corresponding to the extra tape from  $u$  producing a configuration  $u_1$  of  $M$  and check whether  $u_1$  is accepted, i.e. whether  $u_1$  is equivalent to the stop word of  $M$ . This can be done by asking the oracle once. Thus to check whether  $w$  and  $w'$  are equivalent we only need polynomial time and asking the oracle about equivalence of two pairs of configurations, the lengths of which are bounded by  $|w| + |w'|$ . Thus  $M''$  polynomially reduces to  $M$ .  $\square$

**2.3. Minsky machines.** The hardware of a  $K$ -glass Minsky machine,  $K \geq 2$ , consists of  $K$  glasses containing coins. We assume that these glasses are of infinite height. The machine can add a coin to a glass, and remove a coin from a glass (provided the glass is not empty). The commands of a Minsky machine are numbered. So a configuration of a  $K$ -glass Minsky machine is a  $K + 1$ -tuple  $(i; \epsilon_1, \dots, \epsilon_K)$  where  $i$  is the number of command that is to be executed,  $\epsilon_j$  is the number of coins in the glass  $\#j$ .

More precisely, a *command* has one of the following forms:

- Put a coin in each of the glasses  $\#n_1, \dots, n_l$  and go to command  $\#j$ . We shall encode this command as

$$i; \rightarrow \text{Add}(n_1, \dots, n_l); j$$

where  $i$  is the number of the command;

- If the glasses  $\#n_1, \dots, n_l$  are not empty then take a coin from each of these glasses and go to instruction  $\#j$ . This command is encoded as

$$i; \epsilon_{n_1} > 0, \dots, \epsilon_{n_l} > 0 \rightarrow \text{Sub}(n_1, \dots, n_l); j;$$

- If glasses  $\#n_1, \dots, n_l$  are empty, then go to instruction  $\#j$ . This command is encoded as

$$i; \epsilon_{n_1} = 0, \dots, \epsilon_{n_l} = 0 \rightarrow j;$$

- Stop. This command is encoded as  $i; \rightarrow 0$ ;

*Remark 2.5.* This defines deterministic Minsky machines. We will also need non-deterministic Minsky machines. Those will have two or more commands with the same number.

**Theorem 2.6.** *Let  $X$  be a recursively enumerable set of natural numbers. Then the following holds:*

- there exists a 2-glass deterministic Minsky machine  $MM_2$  which recognizes  $L$  in the following sense:  $MM_2$  begins its work in configuration  $(1; 2^m, 0)$  and stops in configuration  $(0; 0, 0)$  if and only if  $m \in X$ , and it works forever if  $m \notin X$ .*
- There exists a 3-glass Minsky machine  $MM_3$  which when started on a configuration  $(1; m, 0, \dots, 0)$  stops in the configuration  $(q_0, 0, 0, \dots, 0)$  provided  $m \in X$ , and works forever otherwise.*
- We can also assume that every computation of  $MM_2$  or  $MM_3$  starting with a configuration  $c$  empties each glass after at most  $O(|c|)$  steps.*
- If  $X$  is recursive, then the machine  $MM_3$  above can be chosen to be sym-universally halting.*
- If  $M$  is a deterministic Turing machine recognizing  $X$ , then we can assume that  $MM_2$  (resp.  $MM_3$ ) polynomially reduces to  $M$  where the numbers written on the tapes of  $M$  are measured as represented in unary (that is the size of a number  $n$  is set as  $n$  and not  $\log_2 n$ ).*

*Proof.* The proof of the 2-glass part can be found in [36]. Let us prove the 3-glass part of the theorem. Let  $M$  be a one tape deterministic Turing machine  $M$  recognizing  $L$ . Without loss of generality we can assume that the tape alphabet of  $M$  is  $\{1, 2\}$ . For every configuration  $\alpha u q_i v \omega$  of  $M$  we can view  $u$  and  $v$  as numbers written in 3-ary, where  $v$  is read from right to left. Let us denote these numbers by  $l(u), r(v)v$ . For example the numbers corresponding to the configuration  $\alpha 121221 q_5 1222 \omega$  are  $l(u) = 121221_3$  and  $r(v) = 2221_3$ , both written as 3-ary numbers. Thus with the configuration  $\alpha u q_i v \omega$ , we associate the following configuration of a 3-tape Minsky machine  $(i; l(u), r(v), 0)$ . Now every command of a Turing machine can be simulated by a series of commands of the Minsky machine. For example, the command  $\Theta$  of the form  $1 q_i 2 \rightarrow q_j 1$  is interpreted by a sequence  $M(\theta)$  of commands of  $MM_3$  as follows. The commands of  $M(\theta)$  will be numbered  $i.1$  through  $i.l$  for some  $l$ . The commands from  $M(\theta)$  should replace  $l(u)$  coins in the first glass by  $\lfloor l(u)/3 \rfloor$  coins provided  $l(u) \equiv 1 \pmod 3$  and replace  $r(v)$  coins in the second glass by  $3 \lfloor r(u)/3 \rfloor + 1$  coins provided  $r(u) \equiv 2 \pmod 3$ . The first part is done as follows. Decrease the number of coins in the first glass by 3 simultaneously increasing the number on the third glass by 1. Do that until the number of coins in the first glass is less than 3. If that remaining number is 1, then subtract 1 coin from the first glass, and then keep adding 1 coin to the first glass, removing 1 coin from the third glass until the third glass is empty. If the remainder is not 1, then keep removing one coin from the third glass while adding 3 coins to the first glass - until the third glass is empty (i.e. return to the original configuration because the command  $\theta$  is not applicable). The second part is done in a similar manner by using the second and third glasses of the Minsky machines. Other commands of the Turing machine are treated in the same manner. Let  $MM_3$  be the resulting 3-tape Minsky machine. It is easy to see that if the Turing machine  $M$  is sym-universally halting, then the Minsky machine  $MM_3$  is sym-universally halting. This gives properties (a),(b) and (d) of the theorem.

To ensure Property (c), we can do the following. Note that after every series of commands  $M(\theta)$  the configuration of  $MM_2$  or  $MM_3$  has (at least) one empty glass. After the series of commands  $M(\theta)$  of  $MM_2$  or  $MM_3$  corresponding to a command  $\theta$  of  $M$  is executed, that glass is again empty. So before  $MM_2$  or  $MM_3$  execute the next series  $M(\theta')$  we force it to move all coins from each of the non-empty glasses to the empty one and back. In the process, it will empty each glass at least once. Clearly, this modification increases the length of computation by an amount proportional to the length of configuration of  $MM_2$  or  $MM_3$ .

Finally property (e) is obtained as follows. Suppose that  $w, w'$  are two configurations of  $MM_3$  (for  $MM_2$  the proof is similar). By construction (see [36]) in at most  $O(|w|)$  steps of  $MM_3$  either  $w$  turns into a configuration corresponding to a configuration of the Turing machine  $M$  or  $MM_3$  halts. In the latter case, we check whether  $w$  is equivalent to  $w'$  in  $O(|w|)$  steps. So we can assume that both  $w$  and  $w'$  are equivalent to configurations corresponding to configurations  $u, u'$  of the Turing machine  $M$  whose lengths are  $O(|w| + |w'|)$ . Now  $w$  is equivalent to  $w'$  if and only if  $u$  and  $u'$  are equivalent configurations of  $M$ . Thus we need to use the oracle once.  $\square$

### 3. SIMULATION OF MINSKY MACHINES BY SEMIGROUPS

**3.1. The construction.** Here we will show how to simulate a Minsky machine by a semi-group. All applications of Minsky machines are based on the following idea.

First, with every configuration  $\psi$  one associates a word (term)  $w(\psi)$ .

Then with every command  $\kappa$  of the Minsky machine  $M$  one associates a finite set of defining relations  $R_\kappa$ . The algebraic structure  $A(M)$  will be defined by the relations from the union  $R$  of all  $R_\kappa$  (which is finite since we have only a finite number of commands) and usually some other relations  $Q$  which are in a sense “independent” of  $R$ . We need  $Q$ , for example, to make sure  $A(M)$  satisfies a particular identity.

We say that the algebra  $A(M)$  *simulates*  $M$  if the following holds for arbitrary configurations  $\psi_1, \psi_2$  of  $M$ :

$$(1) \quad \psi_1 \equiv_M \psi_2 \text{ if and only if } w(\psi_1) = w(\psi_2) \text{ in } A(M).$$

Usually, in order to prove the property (1) one has to prove the following two lemmas.

**Lemma 3.1.** *If a configuration  $c'$  can be obtained from a configuration  $c$  by a command  $\kappa$  of  $M$  then the word  $w(c')$  can be obtained from the word  $w(c)$  by applying defining relations of  $A(M)$  from the set  $R_\kappa$ .*

**Lemma 3.2.** *If a word  $w(c')$  can be obtained from a word  $w(c)$  by applying the defining relations of  $A(M)$  then  $c \equiv_M c'$ .*

It is easy to see that Lemmas 3.1 and 3.2 imply property (1).

There is an easy way to interpret Minsky machines in a semigroup  $S(M)$ . Let  $M$  be a Minsky machine with  $K$  glasses and commands  $\#\#1, 2, \dots, N, 0$ . Then  $S(M)$  is generated by the elements  $q_0, \dots, q_N$  and  $\{a_i, A_i, i = 1, \dots, K\}$ . The set of defining relations of  $S(M)$  consists of all commutativity relations

$$(2) \quad a_i a_j = a_j a_i, a_i A_j = A_j a_i, A_i A_j = A_j A_i, i \neq j,$$

which we shall call *commutativity relations*, the *stop relation*

$$(3) \quad q_0 = 0 \text{ (i.e. } q_0 x = x q_0 = q_0 \text{ for every generator } x),$$

all relations of the form  $xy = 0$  where  $xy$  is a two-letter word which is *not* a subword of a word of the form  $q_i a_1^{\epsilon_1} \dots a_K^{\epsilon_K} A_1 \dots A_K$  modulo the commutativity relations (2), (for example  $q_i q_j = A_i a_i = a_i q_j = A_i q_j = 0$ ), which we shall call *0-relations*, and relations associated with commands of  $M$  according to the following table,

Command of $M$	Relation of $S(M)$
$i \rightarrow \text{Add}(n_1, \dots, n_m); j$	$q_i = q_j a_{n_1} \dots a_{n_m}$
$i, \epsilon_{n_1} > 0, \dots, \epsilon_{n_m} > 0 \rightarrow \text{Sub}(n_1, \dots, n_m); j$	$q_i a_{n_1} \dots a_{n_m} = q_j$
$i, \epsilon_{n_1} = 0, \dots, \epsilon_{n_m} = 0 \rightarrow j$	$q_i A_{n_1} \dots A_{n_m} = q_j A_{n_1} \dots A_{n_m}$

These will be called the Minsky relations.

The words in  $S(M)$  corresponding to configurations of  $M$  are the following:

$$w(i; \epsilon_1, \dots, \epsilon_K) = q_i a_1^{\epsilon_1} \dots a_K^{\epsilon_K} A_1 \dots A_K.$$

The proof that Lemmas 3.1 and 3.2 hold in  $S(M)$  follows easily from Lemma 2.1, see [55, 32].



**Lemma 3.3.** *Suppose that a word  $W$  is not 0 in  $S(M)$ , is a subword of  $w(i; \epsilon_1, \dots, \epsilon_K)$  (up to the commutativity relations (2)), and does not contain either  $q_i$  or one of the  $A_j$ . Then there are at most  $O(|W|)$  different (up to the commutativity relations) words that are equal to  $W$  in  $S(M)$ . All these words are subwords of words of the form  $w(i'; \epsilon_1, \dots, \epsilon_K)$  such that the configurations  $(i; \epsilon_1, \dots, \epsilon_K)$  and  $(i', \epsilon'_1, \dots, \epsilon'_K)$  of  $M$  are equivalent.*

*Proof.* Since  $W \neq 0$  in  $S(M)$ , the stop relations do not apply to  $W$  or to any word that is equal to  $W$  in  $S(M)$ . If  $W$  does not contain  $q_i$ , then the only relations that apply to  $W$  are the commutativity relations, so the only words that are equal to  $W$  in  $S(M)$  are the words obtained from  $W$  by the use of commutativity relations.

Suppose that  $W$  contains  $q_i$  but does not contain one of the  $A_j$ .

Without loss of generality, we can assume that  $W$  contains every letter from  $w(i; \epsilon_1, \dots, \epsilon_K)$  except some of the  $A_j$ 's.

Every application of the Minsky relation to  $W$  corresponds to a command of the Minsky machine, applied to the configuration  $c = (i; \epsilon_1, \dots, \epsilon_K)$ . Let  $c = c_1 \rightarrow c_2 \rightarrow \dots$  be any computation of  $\text{Sym}(M)$  starting with  $c$ . Then the sequence of commands of  $M$  applied in that computation has the form  $\theta_1 \dots \theta_n \theta_{n+1}^{-1} \dots \theta_k^{-1}$  where  $\theta_s$  are commands of  $M$  (by Lemma 2.1). If this sequence can be applied to  $W$ , then this computation never checks whether glass  $\#j$  is empty. By Property (c) of Theorem 2.6, both  $n$  and  $k$  must be at most  $O(|W|)$ . This implies the statement of the lemma.  $\square$

### 3.2. Residually finite finitely presented semigroups.

**Lemma 3.4.** *Every non-zero element of  $S(M)$  is represented by a subword of a word of the form  $w(i; \epsilon_1, \dots, \epsilon_K)$ .*

*Proof.* This follows from the commutativity relations and 0-relations.  $\square$

**Lemma 3.5.** *Suppose that the Minsky machine  $M$  is sym-universally halting. Then*

(a) *Every non-zero element  $z$  of  $S(M)$  has finitely many divisors, i.e. elements  $y$  such that  $z = pyq$  for some  $p, q \in S(M) \cup \{1\}$ .*

(b) *For every configuration  $\psi$  of  $M$  the word  $w(\psi)$  is equal to 0 in  $S(M)$  if and only if  $\psi$  is accepted by  $M$ .*

*Proof.* (a) This follows from Lemmas 2.1, 3.1 and 3.2 for words containing a  $q$ -letter and all letters  $A_j$ , and from Lemma 3.3 in all other cases since word in the generators of  $S(M)$  that is non-zero in  $S(M)$  is a subword of one of the words corresponding to configurations of the Minsky machine  $M$  by Lemma 3.4.

(b) This follows from Lemmas 3.1 and 3.2.  $\square$

Lemma 3.5 immediately implies

**Lemma 3.6.** *For every  $R > 0$  let  $V_R$  be the set of all elements of  $S(M)$  that do not divide in  $S(M)$  non-zero elements represented by words of the form  $q_j a_1^{\epsilon_1} \dots a_K^{\epsilon_K} A_1 \dots A_K$  with  $\epsilon_j \leq R$ . Then  $V_R$  is an ideal of  $S(M)$  with a finite complement. If  $M$  is sym-universally halting, then the intersection of all  $V_R, R > 0$ , is  $\{0\}$ .*

**Theorem 3.7.** *For every recursive set of natural numbers  $Z$  there exists a residually finite semigroup  $S$  whose word problem is at least as hard as the membership problem in  $Z$ . The Dehn function of  $S$  is equivalent to the time function of a 3-glass Minsky machine recognizing  $Z$ .*

*Proof.* By Theorem 2.4, there exists a sym-universally halting Turing machine that recognizes  $Z$ . By Theorem 2.6 there exists a sym-universally halting Minsky machine  $M$  recognizing  $Z$ . By Lemma 3.5, the problem of recognizing equality to 0 in  $S(M)$  is at least as hard as the membership problem in  $Z$ . By Lemma 3.6,  $S(M)$  is residually finite.  $\square$

**3.3. Residually finite semigroups with large depth function.** Recall the definition of the depth function  $\rho$ : for every finitely generated residually finite universal algebra  $A$  and every number  $n$ ,  $\rho_A(n)$  is defined as the smallest number such that for every two different elements  $z, z'$  in  $A$  of length  $\leq n$  there exists a homomorphism  $\phi$  from  $A$  onto a finite algebra  $B$  of cardinality at most  $\rho_A(n)$  such that  $\phi(z) \neq \phi(z')$ .

The following lemma is well known [22]

**Lemma 3.8.** *Suppose that every non-zero element of a semigroup  $S$  with 0 has finitely many divisors. Then  $S$  is residually finite.*

*Proof.* Indeed, the set of all non-divisors of a non-zero element is an ideal with finite quotient. The intersection of all these ideals is  $\{0\}$ .  $\square$

**Theorem 3.9.** *For every recursive function  $f$  there exists a finitely presented residually finite semigroup  $S$  such that  $\rho_S(n) > f(n)$  for all  $n$ . In addition, we can assume that the word problem in  $S$  is as hard as the membership problem for any prescribed recursive set of natural numbers.*

*Proof.* Let  $M$  be a sym-universally halting Minsky machine with  $K$  glasses and  $N + 1$  commands numbered  $0, \dots, N$ . Consider the following new, non-deterministic Minsky machine  $M_n$ . Its hardware consists of the  $K$  glasses of  $M$  plus two more glasses. In every command of  $M$  we add the instruction to add a coin to glass  $K + 1$  provided glass  $K + 2$  is empty. Also for every  $i = 0, \dots, N$  we add two new commands number  $i$

$$(5) \quad i; \text{Add}(K + 1, K + 2) \rightarrow i$$

and

$$(6) \quad i; \epsilon_{K+1} = 0, \epsilon_{K+2} = 0, \rightarrow 0$$

Thus there will be three commands for each  $i = 1, \dots, N$ : one from  $M$  and the two new ones. The new command (5) allows us to add, at any step of the computation, equal (but arbitrary) number of coins in glasses  $K + 1$  and  $K + 2$ , and if both glasses  $K + 1$  and  $K + 2$  are empty, the computation can stop. But we can execute a command of  $M$  only when the glass  $K + 2$  is empty, so a new command cannot be followed by a command of  $M$ .

Let us say that the commands coming from  $M$  have weight 1 and new commands (5), (6) have weight 0. The weight of a computation is then the sum of the weights of all commands used in the computation. We also define the weight of a configuration as the number of coins in the first  $K + 1$  glasses minus the number of coins in glass  $K + 2$ . Every computation  $C$  of  $\text{Sym}(M_n)$  projects onto a computation  $\pi(C)$  of  $\text{Sym}(M)$ : we simply forget the extra two glasses and the new commands. The weight of  $C$  is equal to the length of  $\pi(C)$ . The numbers of coins used in  $C$  and  $\pi(C)$  in the first  $K$  glasses are the same, the number of coins in glass  $K + 1$  in the last configuration  $W$  of  $C$  minus the number of coins in glass  $K + 2$  of  $W$  is equal to the weight of  $C$ .

Also any computation  $C$  of  $M$  lifts to (possibly infinitely many) computations  $C_n$  of  $M_n$ , the weight of each  $C_n$  is the same as the length of  $C$ , and the number of coins used in the first  $K$  glasses is the same.

Note that Lemma 2.1 still holds for  $M_n$  even though  $M_n$  is non-deterministic. It can be easily established by using the projection  $\pi$ .

This implies that if  $M$  is sym-universally halting, then for every configuration  $W$  of  $M_n$  the weights of all computations  $C$  without repeated configurations of  $\text{Sym}(M_n)$  are bounded, the number of coins in the first  $K$  glasses of  $M_n$  used during any of these computations is bounded, and the weights of configurations appearing in these computations are bounded.

Consider the semigroup  $S(M_n)$ . Every non-zero element  $w$  in  $S(M_n)$  is represented by a word of the form  $u(w)v(w)$  where

$$u(w) = q_i^{\alpha_0} a_1^{l_1} \dots a_K^{l_K} A_1^{\alpha_1} \dots A_K^{\alpha_K}, v(w) = a_{K+1}^{l_{K+1}} a_{K+2}^{l_{K+2}} A_{K+1}^{\alpha_{K+1}} A_{K+2}^{\alpha_{K+2}}$$

where  $\alpha_j \in \{0, 1\}, l_j \geq 0$ . Note that if two non-zero words  $w, w'$  are equal in  $S(M_n)$ , then  $u(w)$  and  $u(w')$  are equal in  $S(M)$ .

We claim that  $S(M_n)$  is residually finite. Indeed, consider two words  $w_1, w_2$  in the generators of  $S(M_n)$  which are not equal in  $S(M_n)$ ,  $w_2$  does not divide  $w_1$  in  $S(M_n)$  (clearly  $w_1$  and  $w_2$  cannot divide each other without being equal in  $S(M_n)$ ).

Suppose first that  $w_1$  does not contain a  $q$ -letter. Then consider the ideal  $Q$  of  $S(M_n)$  generated by all  $q$ -letters. The inequality  $w_1 \neq w_2$  survives in the Rees factor-semigroup  $S(M_n)/Q$ . But in  $S(M_n)/Q$  every element has finitely many divisors, hence  $S(M_n)/Q$  is residually finite by Lemma 3.8, and so we can separate  $w_1$  and  $w_2$  by a homomorphism onto a finite semigroup.

Thus we can assume that  $w_1$  starts with a  $q$ -letter  $q_i$ . Suppose that  $u(w_1) \neq u(w_2)$  in  $S(M)$ . Adding the relation  $a_{K+1}^2 = a_{K+1}, a_{K+2}^2 = a_{K+2}^2$  to  $S(M_n)$  we then obtain a new semigroup  $\tilde{S}(M_n)$  and a homomorphism  $\phi: S(M_n) \rightarrow \tilde{S}(M_n)$  which separates  $w_1$  and  $w_2$ . In the semigroup  $\tilde{S}(M_n)$ , every non-zero element has finitely many divisors since it is true for  $S(M)$  and the number of different elements of the form  $v(w)$  is finite. Hence  $\tilde{S}(M_n)$  is residually finite by Lemma 3.8.

Thus we can assume that  $u(w_1) = u(w_2)$ . Let

$$v(w_1) = a_{K+1}^{m_1} a_{K+2}^{m_2} A_{K+1}^{\beta_{K+1}} A_{K+2}^{\beta_{K+2}}, D = \max\{|l_{K+1} - l_{K+2}| + 1, |m_{K+1} - m_{K+2}| + 1\}.$$

Let us add the relations  $a_{K+1}^D = a_{K+1}^{2D}, a_{K+2}^D = a_{K+2}^{2D}$  to  $S(M_n)$ . Let  $\tilde{\tilde{S}}(M_n)$  be the resulting semigroup, and  $\psi: S(M_n) \rightarrow \tilde{\tilde{S}}(M_n)$  be the corresponding homomorphism. Then it is easy to see that  $\psi(w_1) \neq \psi(w_2)$ . Since in  $\tilde{\tilde{S}}(M_n)$ , every element has finite number of divisors (the same argument as for  $\tilde{S}(M_n)$ ), we can again use Lemma 3.8.

The function  $\rho(n)$  for the semigroup  $S(M_n)$  is at least as large as the following function  $\Psi(n)$  associated with the machine  $M$ :  $\Psi(n)$  is the smallest number such that for every non-accepted input configuration of  $M$  of length  $\leq n$ , the machine  $M$  halts after at most  $\Psi(n)$  steps (i.e. the *co-time function* of  $M$ ). Indeed let  $c$  be an input configuration of length at most  $n$  such that  $M$  halts after exactly  $\Psi(n)$  steps starting at  $c$ . Suppose that the word  $w(c)$  in  $S(M_n)$  corresponding to the configuration  $u$  can be separated from 0 in a homomorphic image  $E$  of  $S(M_n)$  with at most  $\Psi(n) - 1$  elements. Then the images of  $a_{K+1}, a_{K+2}$  in that semigroup satisfy  $z^D = z^{2D}$  for some  $D < T(n)$ . Since the halting computation has  $> D$  steps, the letter  $a_{K+1}$  occurs in  $w(u)$  exactly once, and every command of  $M_n$  corresponding to a command of  $M$  adds one coin in glass  $K + 1$ , there exists a word  $W$  which is equal to

$w(u)$  in  $S(M_n)$  and which has the form

$$q_j a_1^{l_1} \dots a_k^{l_K} A_1 \dots A_K a_{K+1}^D A_{K+1} A_{K+2}.$$

Modulo relations corresponding to the commands (5), this word is equal to

$$q_j a_1^{l_1} \dots a_k^{l_K} A_1 \dots A_K a_{K+1}^{2D} A_{K+1} a_{K+2}^D A_{K+2}.$$

The image of the latter word in  $E$  is equal to

$$q_j a_1^{l_1} \dots a_k^{l_K} A_1 \dots A_K a_{K+1}^D A_{K+1} a_{K+2}^D A_{K+2}$$

which, again modulo the relations corresponding to the commands (5), is equal to

$$q_j a_1^{l_1} \dots a_k^{l_K} A_1 \dots A_K A_{K+1} A_{K+2}$$

which is equal to 0 by the relations corresponding to the commands (6), a contradiction.

Note that the co-time function of a Turing machine recognizing a recursive set can be larger than any given recursive function. Indeed, after the machine halts without accepting, we can make it work as long as we like. It remain to note that the co-time function of a Minsky machine simulating that Turing machine cannot be smaller.  $\square$

#### 4. SIMULATION OF MINSKY MACHINES IN SOLVABLE GROUPS

Recall that a *variety* of algebraic structures is a class of all algebraic structures of a given type (signature) satisfying a given set of *identities* (also called *laws*). Equivalently, by a theorem of Birkhoff [36] a variety is a class of algebraic structures closed under taking cartesian products, homomorphic images and substructures. Every variety contains free objects (called *relatively free* algebraic structures). One can define algebraic structures that are finitely presented in a variety as factor-structures by congruence relations generated by finite number of equalities. Every finitely presented algebraic structure which belongs to a variety  $\mathcal{V}$  is finitely presented inside  $\mathcal{V}$  but the converse is very rarely true. See [32] for a survey of algorithmic problems for varieties of different algebraic structures (mostly semigroups, groups, associative and Lie algebras). In this section we concentrate on varieties of groups. The most well known varieties are the variety of Abelian groups  $\mathcal{A}$  given by the identity  $[x, y] = 1$ , the variety of nilpotent groups of class  $c$ ,  $\mathcal{N}_c$  given by the identity  $[\dots[x_1, x_2], \dots, x_{c+1}] = 1$ , etc. The class of Abelian groups of finite exponent  $d$ ,  $\mathcal{A}_d$ , is also a variety, given by two identities  $[x, y] = 1, x^d = 1$ .

If  $\mathcal{U}$  and  $\mathcal{V}$  are two varieties of groups then the class of groups consisting of extensions of groups from  $\mathcal{U}$  by groups from  $\mathcal{V}$  is again a variety (the *product* of  $\mathcal{U}$  and  $\mathcal{V}$ ) denoted by  $\mathcal{UV}$ . The product of varieties is associative [45]. For example the variety of all solvable groups of class  $c$  is the product of  $c$  copies of the variety  $\mathcal{A}$ . If  $\mathcal{V}$  is a variety of groups, then  $\mathcal{ZV}$  is the variety consisting of all central extensions of groups from  $\mathcal{V}$ . For example  $\mathcal{N}_2 = \mathcal{ZA}$  and, more generally,  $\mathcal{N}_{c+1} = \mathcal{ZN}_c$  for every  $c \geq 1$ .

The problem of finding a finitely presented group with undecidable word problem, belonging to a proper variety of groups (i.e. satisfying a non-trivial identity) was formulated by Adian [34] and solved by the first author in [28]. The construction was simplified in the unpublished dissertation [29]. In this section, we shall modify the construction from [29] to construct residually finite finitely presented solvable groups with complicated word problem.

**4.1. The construction.** Let  $M$  be a Minsky machine with  $K$  glasses and  $N+1$  commands (numbered  $0, \dots, N$ ). We are going to construct a group  $G(M)$  simulating  $M$ . The group  $G(M)$  will be in a sense similar to the semigroup  $S(M)$  constructed above. The main idea will be to replace the product by another operation and make sure that with respect to the new operation the semigroup  $S(M)$  “embeds” into our group.

Thus the group will be generated by the  $q$ -letters which will be related to the letters  $q_i$  from  $S(M)$ , and also  $a$ -letters  $a_1, \dots, a_K$ ,  $A$ -letters,  $A_1, \dots, A_K$  and some extra  $a$ - and  $A$ -letters that help us impose the necessary commutativity relations that, in particular, make the group solvable. The group we are going to construct will be a semidirect product of the Abelian normal subgroup generated by the  $q$ -letters by the semidirect product of an Abelian subgroup generated by  $A$ -letters and an Abelian subgroup generated by  $a$ -letters. Thus we should have a way to ensure that in a subgroup generated by two sets of letters  $Z \cup Y$ , the normal subgroup generated by  $Z$  is Abelian. This is done with the help of the following lemma due to Baumslag [3] and Remeslennikov [52]. In that lemma we denote  $u^a = a^{-1}ua$  and  $u^{a+b} = u^a u^b$  (note that although  $u^{a+b}$  is not necessarily equal to  $u^{b+a}$ , the equality will hold if the normal subgroup generated by  $u$  is Abelian, which is going to be the case every time we apply this lemma).

**Lemma 4.1** ([3, 52]). *Suppose that a group  $H$  is generated by three sets  $X, F = \{a_i \mid i = 1, \dots, m\}, F' = \{a'_i \mid i = 1, \dots, m\}$  such that*

- (1) *The subgroup generated by  $F \cup F'$  is Abelian;*
- (2) *For every  $a \in F$  and every  $x \in X$  we have  $x^{f(a)} = x^{a'}$  for some monic polynomial  $f$  of  $a$  which has at least two terms (in all our applications  $f(t) = t - 1$ );*
- (3)  *$[x_1^{a_1^{\alpha_1} \dots a_m^{\alpha_m}}, x_2] = 1$ , for every  $x_1, x_2 \in X$ , and every  $\alpha_1, \dots, \alpha_m \in \{0, 1, -1\}$ .*

*Then the normal subgroup generated by  $X$  in the group  $H = \langle X \cup F \cup F' \rangle$  is Abelian, and  $H$  is metabelian.*

If the elements  $a_i$  and  $a'_i$  and the set  $X$  satisfy the conditions of Lemma 4.1 we will call  $a'_i, i = 1, \dots, m$ , are BR-conjoints to  $a_i$  with respect to  $X$  (and the polynomial  $f$ ).

Consider the free commutative monoid generated by letters  $A_0, \dots, A_K$ . Let  $U_0$  be the set of all divisors of the element  $A_0 A_1 \dots A_K$  in that monoid, and  $U$  be the set of all symbols  $q_j w, w \in U_0, j = 0, \dots, N$ . Also fix a prime  $p$  (say,  $p = 2$ ).

The generating set of our group  $G = G(M)$  will consists of three subsets:

$$\begin{aligned} L_0 &= \{x_u, u \in U, i = 0, \dots, N\}; \\ L_1 &= \{A_i, i = 0, \dots, K, \}; \\ L_2 &= \{a_i, a'_i, \tilde{a}_i, \tilde{a}'_i, i = 1, \dots, K\}. \end{aligned}$$

We introduce notation for some subgroups of the group  $G$ . Denote  $H_i = \langle L_i \rangle, i = 0, 1, 2$ . Denote also

$$M_0 = \{\tilde{a}_i, \tilde{a}'_i, A_0, i = 1, \dots, K\}, \quad M_i = \{a_i, a'_i, A_i\}, i = 1, \dots, K$$

The group  $G(M)$  has the following set of defining relations:

G1. Relations saying that  $H_0$  and  $H_1$  are Abelian groups of exponent  $p$ , and  $H_2$  is an Abelian group.

G2. Any  $y \in M_i, z \in M_j, i \neq j \in \{0, \dots, K\}$ , commute.

G3. For every  $i = 1, \dots, K, (a'_i)^{-1}$  is a BR-conjoint to  $a_i^{-1}$  with respect to  $\{A_i\}$  (and polynomial  $f(t) = t - 1$ ).

G4. The elements of the set  $\{(\tilde{a}'_i)^{-1}, i = 1, \dots, K\}$  are a BR-conjoints to elements of the set  $\{\tilde{a}_i^{-1}, i = 1, \dots, K\}$  with respect to  $\{A_0\}$ .

G5. a) If  $u \in U$  does not contain  $A_i$  for some  $i = 0, \dots, K$ , then  $[x_u, A_i] = x_{uA_i}$ ,  $j = 0, \dots, K$ .

b) For every  $i = 1, \dots, K$ , if  $u$  does not contain  $A_i$ , then  $x_u^{a_i-1} = x_u^{a'_i}$  (see notation before Lemma 4.1,

c) For every  $i = 0, \dots, K$ , if  $u$  contains  $A_i$ ,  $z \in M_i$ , then  $[x_u, z] = 1$ .

G6.  $x_{q_j}^{a_i} = x_{q_j}^{\tilde{a}_i}$ ,  $x_{q_i}^{a'_i} = x_{q_i}^{\tilde{a}'_i}$ ,  $j = 0, \dots, N$ ,  $i = 1, \dots, K$ .

G7.  $[x_u^z, x_v] = 1$ , where  $z = a_1^{\alpha_1} \dots a_K^{\alpha_K}$ ,  $\alpha_i \in \{-1, 0, 1\}$ .

*Remark 4.2.* Relations G7 together with G1 and G5b) imply that for every subset  $I \subseteq \{1, \dots, K\}$  the letters  $\{a'_i, i \in I\}$  are BR-conjoints of  $\{a_i, i \in I\}$  with respect to the set of all  $x_u$ 's where  $u$  does not contain letters  $A_i, i \in I$ .

G8. Relations constructed from the program of the machine  $M$ . For every  $f \in G$  denote

$$f * a_i = f^{-1} f^{a_i} f^{-a_i^{-1}} f^{(a'_i)^{-1}}, i = 1, \dots, K,$$

also let

$$f * A_i = [f, A_i], i = 0, \dots, K.$$

We denote  $(\dots(t_1 * t_2) * \dots) * t_k$  by  $t_1 * \dots * t_k$ , and  $t_1 * \underbrace{t_2 * \dots * t_2}_n$  by  $t_1 * t_2^{(n)}$ . The relations corresponding to the commands of  $M$  are in the following table.

Command of $M$	Relation of $G(M)$
$i \rightarrow \text{Add}(n_1, \dots, n_m); j$	$x_{q_i A_0} = x_{q_j A_0} * a_{n_1} * \dots * a_{n_m}$
$i, \epsilon_{n_1} > 0, \dots, \epsilon_{n_m} > 0 \rightarrow \text{Sub}(n_1, \dots, n_m); j$	$x_{q_i A_0} * a_{n_1} * \dots * a_{n_m} = x_{q_j A_0}$
$i, \epsilon_{n_1} = 0, \dots, \epsilon_{n_m} = 0 \rightarrow j$	$x_{q_i A_0} * A_{n_1} * \dots * A_{n_m} = x_{q_j A_0} * A_{n_1} * \dots * A_{n_m}$

**Theorem 4.3.** (a) The group  $G(M)$  belongs to  $\mathcal{A}_p^2 \mathcal{A} \cap \mathcal{ZN}_{K+1} \mathcal{A}$ .

(b) The equality

$$x_{q_i A_0} * a_1^{(m_1)} * \dots * a_K^{(m_K)} * A_1^{(\alpha_1)} * \dots * A_K^{(\alpha_K)} = x_{q_j A_0} * a_1^{(n_1)} * \dots * a_K^{(n_K)} * A_1^{(\beta_1)} * \dots * A_K^{(\beta_K)}$$

where  $\alpha_i, \beta_i \in \{0, 1\}$  is true in  $G(M)$  if and only if the equality

$$q_i a_1^{m_1} \dots a_K^{m_K} A_1^{\alpha_1} \dots A_K^{\alpha_K} = q_j a_1^{n_1} \dots a_K^{n_K} A_1^{\beta_1} \dots A_K^{\beta_K}$$

is true in the semigroup  $S(M)$  (in particular,  $\alpha_i = \beta_i$  for every  $i$ ).

(c) The equality

$$x_{q_i} * a_1^{(m_1)} * \dots * a_K^{(m_K)} * A_1^{(\alpha_1)} * \dots * A_K^{(\alpha_K)} = x_{q_j} * a_1^{(n_1)} * \dots * a_K^{(n_K)} * A_1^{(\beta_1)} * \dots * A_K^{(\beta_K)}$$

where  $\alpha_i, \beta_i \in \{0, 1\}$  is true in  $G(M)$  if and only if  $m_i = n_i, \alpha_i = \beta_i$  for every  $i$ .

*Proof.* First we will prove part (a):  $G(M) \in \mathcal{A}_p^2 \mathcal{A} \cap \mathcal{ZN}_3 \mathcal{A}$ .

**Lemma 4.4.** The subgroup  $\langle H_1 \cup H_2 \rangle$  of  $G$  is metabelian and a semidirect product of the Abelian normal subgroup  $H_1^{H_2}$  of exponent  $p$ , and  $H_2$ .



*Proof.* Indeed by relations G2,

$$\langle M_i, i = 0, \dots, K \rangle = \prod_{i=1}^K \langle a_i, a'_i, A_i \rangle \times \langle \tilde{a}_i, \tilde{a}'_i, A_0, i = 1, \dots, K \rangle.$$

Using relations G1, G3, G4, we can apply Lemma 4.1 to each of the factors in that direct product and conclude that each of them is metabelian and a semidirect product of the Abelian of exponent  $p$  normal subgroup generated by the intersection of  $\{A_i, i = 0, \dots, K\}$  with that factor, and the Abelian group generated by the  $a$ -letters from that factor.  $\square$

**Lemma 4.5.** *The normal subgroup  $T$  of  $G$  generated by all the elements  $x_u, u \in U$ , is Abelian of exponent  $p$ .*

*Proof.* Relations G5 a) of the group  $G$  imply that every element  $x_u, u \in U$  is a product of elements  $x_{q_j}^z, z \in H_1, i = 0, \dots, N$ . Therefore, it is enough to show that

$$(8) \quad x_{q_k} x_{q_t}^z = x_{q_t}^z x_{q_k}$$

for any  $z \in \langle H_1, H_2 \rangle$  and any  $k, t$ . To reduce the proof of these equalities to the proof of more simple equalities notice that  $z = z_0 z_1 \dots z_K$  where  $z_i \in \langle M_i \rangle$  by G2. Therefore equalities (8) are equivalent to

$$(9) \quad x_{q_k}^{z_0} x_{q_t}^{z_1 \dots z_K} = x_{q_t}^{z_1 \dots z_K} x_{q_k}^{z_0}.$$

We can represent element  $x_{q_j}^{z_i}, i \geq 1$ , as a product of elements of the form  $x_{q_j}^{a_i^p (a'_i)^q}$  and  $x_{q_j A_i}^{\tilde{a}_i^p (\tilde{a}'_i)^q}$ . Indeed we have the following sequence of equalities deduced using G2, G5, G6:

$$\begin{aligned} & x_{q_j}^{a_i^{r_1} (a'_i)^{s_1} A_i^{t_1} \dots a_i^{r_s} (a'_i)^{s_k} A_i^{t_k}} \stackrel{\text{G6}}{=} x_{q_j}^{\tilde{a}_i^{r_1} (\tilde{a}'_i)^{s_1} A_i^{t_1} \dots a_i^{r_k} (a'_i)^{s_k} A_i^{t_k}} \\ & \stackrel{\text{G2}}{=} x_{q_j}^{A_i^{s_1} \tilde{a}_i^{r_1} (\tilde{a}'_i)^{s_1} a_i^{r_2} (a'_i)^{s_2} \dots a_i^{r_k} (a'_i)^{s_k} A_i^{t_k}} \\ (10) \quad & \stackrel{\text{G5 a), c), G6}}{=} x_{q_j}^{a_i^{r_1+r_2} (a'_i)^{s_1+s_2} A_i^{t_2} \dots a_i^{r_k} (a'_i)^{s_k} A_i^{t_k}} (x_{q_j A_i}^{t_1})^{\tilde{a}_i^{r_1} (\tilde{a}'_i)^{s_1}} = \\ & \dots = x_{q_j}^{a_i^{r_1+\dots+r_k} (a'_i)^{s_1+\dots+s_k}} (x_{q_t A_i}^{t_k})^{\tilde{a}_i^{r_1+\dots+r_k} (\tilde{a}'_i)^{s_1+\dots+s_k}} \dots (x_{q_t A_i}^{t_1})^{\tilde{a}_i^{r_1} (\tilde{a}'_i)^{s_1}}. \end{aligned}$$

Repeating this argument  $K$  times, one proves that  $x_{q_j}^{z_1 z_2 \dots z_K}$  can be represented as a product of elements of the form  $x_u^y$  where  $u \in U, y \in H_2$ . A similar proof (using also G4) gives that  $x_{q_j}^{z_0}$  is a product of elements of that form. It remains to note that elements of the form  $x_u^y, u \in U, y \in H_2$  commute by Remark 4.2 and Lemma 4.1.  $\square$

*Remark 4.6.* Note that equalities (10) and similar equalities when  $x_{q_j}$  is replaced by  $x_u, u \in U$ , imply the following: if  $y$  is a product of elements of the form  $a_i^{r_l} (a'_i)^{s_l} A_i$  and  $\sum_l r_l = \sum_l s_l = 0$ , then  $[x_u, y]$  is 1 if  $u$  contains  $A_i$  or a product of conjugates of elements  $x_{u A_i}$  by elements from  $\langle \tilde{a}_i \rangle \times \langle \tilde{a}'_i \rangle$  otherwise. Similarly, suppose that  $y$  is a product of elements from  $M_0$ , each factor containing  $A_0$ , and the total exponent of every  $\tilde{a}_i$  (resp.  $\tilde{a}'_i$ ) is 0. Then  $[x_u, y] = 1$  provided  $u$  contains  $A_0$  and is a product of conjugates of  $x_{u A_0}$  by elements from  $\langle a_i, a'_i \rangle$  provided  $u$  does not contain  $A_0$ .

By construction, the group  $G$  is a semidirect product of  $T$  and the metabelian group  $H_1^{H_2} \rtimes H_2$ . By Lemma 4.4,  $G$  is solvable of class 3 and, moreover, belongs to  $\mathcal{A}_p^2 \mathcal{A}$ .

*Remark 4.7.* The proof of Lemma 4.5 shows that  $T$  is generated (as an Abelian group) by elements of the form  $x_u^y$  where  $u \in U$  and  $y \in H_2$ .

**Lemma 4.8.** *The quotient of  $G(M)$  over the center satisfies the identity*

$$[[x_1, y_1], [x_2, y_2], \dots, [x_{K+2}, y_{K+2}]] = 1.$$

*This means that  $G$  belongs to the variety  $\mathcal{ZN}_{K+1}\mathcal{A}$ .*

*Proof.* Let  $P$  be the derived subgroup of  $G(M)$ . By Lemma 4.5, every element of  $P$  is a product of an element of  $T$  and an element of  $H_1^{H_2}$ . It also follows from Lemma 4.5 that  $[P, P] \subseteq T$ , hence by Remark 4.7, it is generated by elements of the form  $x_u^y$ ,  $u \in U, y \in H_2$ , the word  $u$  contains at least one  $A_i$ ,  $i = 0, \dots, K$ . Since  $T$  is Abelian, the subgroup  $\underbrace{[P, P, \dots, P]}_{K+2}$  is generated by the commutators

$$[x_u^y, h_{1,1}^{h_{2,1}}, \dots, h_{1,K}^{h_{2,K}}]$$

for some  $h_{1,i} \in H_1$ ,  $y, h_{2,i} \in H_2$ . An easy induction shows that every such commutator is a conjugate of

$$(11) \quad [x_u, h_{1,1}^{y'}, \dots, h_{1,K}^{y'}]$$

where  $y' \in H_2$ .

Let  $h \in H_1$ ,  $u \in U, y \in H_2$ . Suppose that  $h = A_{i_1}^{t_1} \dots A_{i_s}^{t_s}$  where  $t_i \neq 0$ . Consider  $[x_u, h^y]$ . Then Remark 4.6 implies that  $[x_u, h^y]$  is a product of elements of the form  $x_{u'}^{y'}$  where  $u' \in U$  contains letters  $A_{i_1}, \dots, A_{i_s}$  and it may not be equal to 1 only if one of the letters  $A_{i_j}$  does not occur in  $u$ . Therefore the commutator (11) is either equal to 1 or is a product of elements of the form  $x_{u'}^{y''}$  where the word  $u' \in U$  contains all letters  $A_0, A_1, \dots, A_K$ ,  $y'' \in H_2$ . But every such  $x_{u'}$  is in the center of  $G(M)$  by G5 c). Hence  $\underbrace{[P, \dots, P]}_{K+2}$  is contained in the center of

$G(M)$ . □

We now prove (b) and (c). For this, as we mentioned before Lemma 3.1, we need to prove Lemmas 3.1 and 3.2. Lemma 3.1 for  $G(M)$  is proved in the same way as for the semigroup  $S(M)$  (see [55, 32], since the only property of  $S(M)$  used there was that the word  $w = q_i a_1^{l_1} \dots a_K^{l_K} A_1^{\alpha_1} \dots A_K^{\alpha_K}$  is equal to any word obtained from  $w$  by permuting  $a_i$  with  $a_j$ ,  $A_i$  with  $A_j$  and  $a_i$  with  $A_j$  ( $i \neq j$ ). The same is true for words of the form

$$(12) \quad x_{q_i A_0} * a_1^{(m_1)} * \dots * a_K^{(m_K)} * A_1^{(\alpha_1)} * \dots * A_K^{(\alpha_K)}$$

in  $G(M)$  by the definition of the operation  $*$ , relations G1, G2 and Lemma 4.5.

In order to prove Lemma 3.2 we will define a new group  $\bar{G}$  that is a quotient of  $G$  and injective on elements of the form (12).

Let  $\check{S}$  be the semigroup with the same generating set as  $S(M)$  subject all the relations of  $S(M)$  except the relations (4) corresponding to the commands of  $M$  (that semigroup does not depend on  $M$ ). Thus non-zero elements in  $\check{M}$  have the form

$$q_i^{\alpha_1} a_1^{l_1} \dots a_K^{l_K} A_1^{\alpha_1} \dots A_K^{\alpha_K}$$

where  $l_j \in \mathbb{N}$ ,  $\alpha_j \in \{0, 1\}$ . Let  $W$  be the set of all non-zero elements of  $\check{S}$  containing a  $q$ -letter, and  $W_0$  be the set of elements from  $W$  viewed as elements of  $S(M)$  (i.e. different words may represent equal element) with  $A_0$  inserted next to the  $q$ -letter. Consider the

free Abelian group  $T_1$  of exponent  $p$  generated by the elements  $z_{i_1, \dots, i_K, u}$ ,  $u \in W \cup W_0$ ,  $i_j \in \{1, 2, 3\}$ . For each element of  $L_1 \cup L_2$ , we define an automorphism of  $T_1$ . The group  $\tilde{G}$  will be the semidirect product of  $T_1$  and the group generated by these automorphisms.

For simplicity we will denote automorphisms corresponding to letters from  $L_1 \cup L_2$  by the same letters.

Let us start with automorphisms  $a_j$ ,  $a'_j$ . We have to define  $z_{i_1, \dots, i_K, u}^{a_i}$  and  $z_{i_1, \dots, i_K, u}^{a'_i}$  for every  $i_1, \dots, i_K$ . First suppose that  $u$  does not contain  $A_j$ . To simplify the notation we shall denote the vector  $(i_1, \dots, i_K)$  by  $\vec{i}$ , and the standard unit vectors by  $\vec{e}_l, l = 1, \dots, K$ . We shall write  $z_{\vec{i}, u}$  instead of  $z_{i_1, \dots, i_K, u}$ . The  $j$ -th coordinate of  $\vec{i}$  is denoted by  $\vec{i}_j$ .

$$(13) \quad z_{\vec{i}, u}^{a_j} = \begin{cases} z_{\vec{i}, u} z_{\vec{i} + \vec{e}_j, u} z_{\vec{i} + 2\vec{e}_j, u} z_{\vec{i}, ua} & \text{if } \vec{i}_j = 1; \\ z_{\vec{i}, u} z_{\vec{i} - \vec{e}_j, u}^{-1} & \text{if } \vec{i}_j = 2; \\ z_{\vec{i} - 2\vec{e}_j, u} & \text{if } \vec{i}_j = 3. \end{cases}$$

$$z_{\vec{i}, u}^{a'_j} = z_{\vec{i}, u}^{-1} z_{\vec{i}, u}^{a_j}.$$

If  $u$  contains letter  $A_j$ , then let  $z_{\vec{i}, u}^{a_j} = z_{\vec{i}, u}^{a'_j} = z_{\vec{i}, u}$ .

It is easy to prove that  $a_j$  is an automorphism by constructing the automorphism  $a_j^{-1}$ . If we apply  $a_j^{-1}$  to the third equality in (13), we will obtain the formula for  $z_{\vec{i}, u}^{a_j^{-1}}$  provided  $\vec{i}_j = 1$  (and  $u$  does not contain  $A_j$ ). Plugging it in the second equality of (13) we obtain the formula for  $z_{\vec{i}, u}^{a_j^{-1}}$  provided  $\vec{i}_j = 2$ . Finally plugging it in the first equality in (13), we obtain the formula for  $z_{\vec{i}, u}^{a_j^{-1}}$  provided  $\vec{i}_j = 3$ :

$$x_{\vec{i}, u}^{a_j^{-1}} = \begin{cases} x_{\vec{i} - \vec{e}_j, u}^{-1} x_{\vec{i}, u}^{-1} x_{\vec{i}, ua}^{-1}, & \text{if } \vec{i}_j = 3 \\ x_{\vec{i}, u} x_{\vec{i} + \vec{e}_j, u}, & \text{if } \vec{i}_j = 2 \\ x_{\vec{i} + 2\vec{e}_j, u}, & \text{if } \vec{i}_j = 1. \end{cases}$$

The automorphism  $\tilde{a}_j$  is defined similarly. If  $u$  contains  $A_0$ , then  $z_{\vec{i}, u}^{\tilde{a}_j} = z_{\vec{i}, u}$ . If  $u$  does not contain  $A_0$  and  $A_j$  then  $z_{\vec{i}, u}^{\tilde{a}_j} = z_{\vec{i}, u}^{a_j}$ .

If  $u$  does not contain  $A_0$  but contains  $A_j$ , i.e.  $u = vA_j$  for some  $v$ , then

$$z_{\vec{i}, u}^{\tilde{a}_j} = \begin{cases} z_{\vec{i}, u} z_{\vec{i} + \vec{e}_j, u} z_{\vec{i} + 2\vec{e}_j, u} z_{\vec{i}, vA_j A_j}, & \text{if } \vec{i}_j = 1; \\ z_{\vec{i}, u} z_{\vec{i} - \vec{e}_j, u}^{-1}, & \text{if } \vec{i}_j = 2; \\ z_{\vec{i} - 2\vec{e}_j, u}, & \text{if } \vec{i}_j = 3. \end{cases}$$

$$z_{\vec{i}, u}^{\tilde{a}'_j} = z_{\vec{i}, u}^{-1} z_{\vec{i}, u}^{\tilde{a}_j}.$$

Finally the automorphisms corresponding to  $A_j$ ,  $j = 0, \dots, K$ , are defined as follows:

$$z_{\vec{i}, u}^{A_j} = z_{\vec{i}, u} z_{\vec{i}, uA_j}$$

if  $u$  does not contain  $A_j$  and

$$z_{i,u}^{A_j} = z_{i,u}^{\vec{z}}$$

if  $u$  contains  $A_j$ .

The following lemma is obtained by a straightforward application of the definition of the automorphisms above and the definition of the operation  $*$ . This lemma implies that  $\bar{G}$  satisfies G8 if we replace  $x_u$  by  $z_{\vec{1},u}$  (since the corresponding relations hold in  $S(M)$ ).

**Lemma 4.9.** *The following relations hold in  $\bar{G}$ . For every  $w \in \{a_j, A_j, j = 1, \dots, K\}$*

$$z_{\vec{1},u} * w = z_{\vec{1},uw}$$

where we set  $z_{\vec{1},0} = 1$  (the identity element in  $G(M)$ ) where  $*$  is defined in G8.

We define  $\bar{G}$  as the semidirect product of  $T_1$  and the subgroup of  $\text{Aut}(T_1)$  generated by the automorphisms corresponding to the elements from  $L_1 \cup L_2$ . From the definition of the automorphisms and Lemma 4.9, it follows that  $\bar{G}$  is generated by the elements  $z_{\vec{1},u}$ ,  $u \in U$ , where  $\vec{1}$  is the vector  $(1, 1, \dots, 1)$  and the automorphisms corresponding to elements of  $L_1 \cup L_2$ . It is easy to check that all the relations G1-G8 hold in  $\bar{G}$ , therefore

**Lemma 4.10.** *The map that sends every  $a$ - or  $A$ -letter to itself and every  $x_u$  to  $z_{\vec{1},u}$  extends to a homomorphism  $\phi$  from  $G$  to  $\bar{G}$ .*

**Lemma 4.11.** *The homomorphism  $\phi$  is surjective.*

*Proof.* It is easy to see that we only need to define pre-images  $x_{i,w}^{\vec{z}}$  of elements  $z_{i,u}^{\vec{z}} \in \bar{G}$ ,  $w \in W \cup W_0$ . By the definition of  $\phi$ , we have  $\phi(x_u) = z_{\vec{1},u}$  for every  $u \in U$  so we define  $x_{\vec{1},u} = x_u$ . The other preimages are defined by induction on the length of  $w$  and the sum of  $i_j$ .

Suppose  $w \in W \cup W_0$  does not contain  $A_j$  and  $\vec{i}_j = 1$ ,  $\vec{i}'$  is arbitrary. Then we define:

$$\begin{aligned} x_{\vec{i}+\vec{e}_j,w}^{\vec{z}} &= x_{i,w}^{\vec{z}}^{-(a'_j)^{-1}}, \\ x_{\vec{i}+2\vec{e}_j,w}^{\vec{z}} &= x_{i,w}^{\vec{z}}^{a_j^{-1}}, \\ x_{\vec{i}',w}^{\vec{z}} * a_j &= x_{i',wa_j}^{\vec{z}}. \end{aligned}$$

We also have  $x_{i,w}^{\vec{z}} * A_j = x_{i,wa_j}^{\vec{z}}$  for any  $\vec{i}$ .

It is easy to see that for every  $\vec{i}$  and  $w \in W \cup W_0$ , we have  $\phi(x_{i,w}^{\vec{z}}) = z_{i,w}^{\vec{z}}$ . This proves the lemma.  $\square$

In  $\bar{G}(M)$ , consider the set  $P$  of elements

$$(14) \quad z_{\vec{1},q_i A_0} * a_1^{(m_1)} * \dots * a_K^{(m_K)} * A_1^{(\alpha_1)} * \dots * A_K^{(\alpha_K)}$$

where  $\alpha_i \in \{0, 1\}$  and the set  $P_0$  of elements

$$(15) \quad z_{\vec{1},q_i} * a_1^{(m_1)} * \dots * a_K^{(m_K)} * A_1^{(\alpha_1)} * \dots * A_K^{(\alpha_K)}$$

By construction  $P \cap P_0 = \emptyset$ , elements (14) are different if and only if elements

$$q_i a_1^{m_1} \dots a_K^{m_K} A_1^{\alpha_1} \dots A_K^{\alpha_K}$$

from  $S(M)$  are different, and elements (15) are different if and only if the corresponding elements  $q_i a_1^{m_1} \dots a_K^{m_K} A_1^{\alpha_1} \dots A_K^{\alpha_K}$  of  $\check{S}$  are different. This completes the proof of Lemma 3.2 and Theorem 4.3 (b), (c).  $\square$

We shall need a few more properties of the group  $G(M)$ .

**Lemma 4.12.** *Let elements  $x_{\vec{i},w}$ ,  $w \in W \cup W_0$  from  $G$  be defined as in the proof of Lemma 4.11. Let  $y \in L_1 \cup L_2$ ,  $w \in W \cap W_0$ . Then for every  $i \in \{1, 2, 3\}^{\{1, \dots, K\}}$ ,  $x_{\vec{i},u}^y$  satisfies the same equalities as elements  $z_{\vec{i},w}^y$  from the definition of automorphism of  $\bar{G}$  with  $z$  replaced by  $x$  everywhere. In particular,  $x_{\vec{i},u}^y$  is a product of one or several elements of the form  $x_{\vec{i}',w'}$ , such that every letter  $a_j$  occurs in  $w'$  at least as many times as in  $w$  (in particular if for some  $R > 0$ ,  $w$  belongs to the ideal  $V_R$  defined in Lemma 3.6, then  $w' \in V_R$ ).*

*Proof.* For  $y \in \cup M_j$ ,  $j \geq 1$ , this follows from the way  $x_{\vec{i},u}$  are constructed. For  $y \in M_0$ , one needs to use G2, G5 c), and G6.  $\square$

The proof of Lemma 4.12 actually gives the following

**Lemma 4.13.** *If  $v$  is a word in  $a$ - and  $A$ -letters (i.e. over  $L_1 \cup L_2$ ), then  $x_{\vec{i},u}^v$  is a product in  $G$  of elements  $x_{\vec{j},w}$  as in Lemma 4.12 where the length of each  $w$  does not exceed the length of  $v$  (hence the total number of different  $x_{\vec{j},w}$  occurring in this product is polynomial in terms of  $|v|$ ).*

**Lemma 4.14.** *The normal subgroup  $T$  generated by the elements  $x_u$ ,  $u \in U$  in  $G = G(M)$  is the direct product of cyclic subgroups generated by the elements  $x_{\vec{i},w}$ ,  $\vec{i} \in \{1, 2, 3\}^{\{1, \dots, K\}}$ ,  $w \in W \cup W_0$ .*

*Proof.* By Lemma 4.12 elements  $x_{\vec{i},w}$  span  $T$ . We defined elements  $x_{\vec{i},w}$ ,  $w \in W$  in such a way that they are pre-images of the corresponding elements in  $\bar{G}$  under  $\phi$ . Thus the elements  $x_{\vec{i},w}$ ,  $\vec{i} \in \{1, 2, 3\}^{\{1, \dots, K\}}$ ,  $w \in W \cup W_0$  are linearly independent since their images under  $\phi$  are linearly independent in  $T_1$ .  $\square$

**Lemma 4.15.** *Let  $R > 0$ ,  $V_R$  be the ideal of the semigroup  $S(M)$  defined in Lemma 3.6. Then the subgroup  $T(V_R)$  of  $T$  spanned by all the elements  $x_{\vec{i},w}$ ,  $w \in V_R$ ,  $\vec{i} \in \{1, 2, 3\}^{\{1, \dots, K\}}$  is normal in  $G(M)$  (as before, we set  $x_{\vec{i},0} = 1$ ) and of finite index in  $T$ .*

*Proof.* The first part follows from Lemma 4.12. If  $\{w_1, \dots, w_M\}$  is the set  $S(M) \setminus V_R$ , then  $\{x_{\vec{i},w}, \vec{i} \in \{1, 2, 3\}^{\{1, \dots, K\}}, w \in \{w_1, \dots, w_M\}\}$  is a set of representatives of all cosets of  $T(V_R)$  in  $T$ .  $\square$

**4.2. A finitely presented solvable group with undecidable word problem.** By Theorem 2.6, there exists a 2-glass Minsky machine which computes a non-recursive partial function. The corresponding group  $G(M)$  has undecidable word problem and belongs to the variety  $\mathcal{A}_p^2 \mathcal{A} \cap \mathcal{ZN}_3 \mathcal{A}$  by Theorem 4.3. Hence we obtain the following:

**Theorem 4.16** (Kharlampovich [28]). *There exists a finitely presented group with undecidable word problem that belongs to the variety  $\mathcal{A}_p^2 \mathcal{A} \cap \mathcal{ZN}_3 \mathcal{A}$ .*

### 4.3. Residually finite finitely presented groups.

**Theorem 4.17.** *If a Minsky machine  $M$  is sym-universally halting then the group  $G(M)$  is residually finite. Its word problem is at least as hard as the halting problem for  $M$ .*

*Proof.* Let  $M$  be a sym-universally halting Minsky machine. Let  $w \neq 1 \in G(M)$ . We use the notation from the definition of  $G(M)$ . There exists a natural homomorphism  $\zeta$  from  $G(M)$  to the metabelian group  $H_1^{H_2} \rtimes H_2$  which kills all elements from  $T$ . Since every finitely generated metabelian group is residually finite, we can assume that  $\zeta(w) = 1$ . Hence  $w \in T$ . By Lemma 4.14,  $x$  is a product of elements of the form

$$(16) \quad x_{\vec{i}, u}, \vec{i} \in \{1, 2, 3\}^{\{1, \dots, K\}}, u \in W \cup W_0.$$

Hence  $w = w_0 w_1$  where  $w_0$  (resp.  $w_1$ ) is a product of elements (16) with  $u \in W_0$  (resp.  $W$ ). Suppose that  $w_1$  is not 1. Let  $T'$  be the subgroup of  $G(M)$  generated by elements (16) with  $u \in W_0$ . Then  $T'$  is a normal subgroup of  $G(M)$  by Lemma 4.12. Let  $G'(M) = G(M)/T'$ . This group is a semidirect product of  $T/T'$  and the metabelian group  $H_1^{H_2} \rtimes H_1$ . Let  $D$  be the sum of lengths of words  $u \in W$  that appear in the factors of  $w_1$ . Let  $Y_D$  be the set of all words in  $\check{S}$  where at least one  $a$ -letter appears at least  $D$  times, and 0. Then  $Y_D$  is an ideal in  $\check{S}$ , and the image of the set of elements (16) with  $u \in Y_D$  in  $G'(M)$  form a normal  $N$  subgroup of  $G'(M)$  of finite index (because  $T$  is an Abelian group of finite exponent  $p$ ). That normal subgroup does not contain  $w$  by Theorem 4.3 (c). Then  $G'(M)/N$  is a semidirect product of a finite group and the metabelian group  $H_1^{H_2} \rtimes H_2$ . Hence  $G'(M)/N$  is residually finite and  $w$  can be separated from 1 by a homomorphism from  $G(M)$  onto a finite group.

Finally suppose that  $w_1 = 1$ . Let  $u_1, \dots, u_l$  be the elements from  $W_0$  that appear in the representation of  $w$  as a product of elements (16). Let  $E$  be the set of words that is equal to one of the  $u_j$  in  $S(M)$ . Since  $M$  is sym-universally halting,  $E$  is finite. Let  $D$  be the maximal length of a word in  $E$ . Let, as above,  $Y_D$  be the ideal in  $\check{S}$  consisting of 0 and all elements where one of the  $a$ -letters appears at least  $D$  times. Let  $Z_D$  be the set of non-zero elements of  $S(M)$  that are images of words from  $Y_D$  under the natural homomorphism  $\check{S} \rightarrow S(M)$ . Then  $Z_D$  does not contain  $u_1, \dots, u_l$ . Consider the subgroup  $F$  of  $T$  spanned by all elements (16) with  $u \in Z_D \cup Y_D$ . From Lemma 4.12, it follows that  $F$  is a normal subgroup of  $G(M)$  of finite index in  $T$ . Since  $Z_D$  does not contain  $u_1, \dots, u_l$ , the subgroup  $F$  does not contain  $w$ . The factor-group  $G(M)/F$  is a semidirect product of a finite group and the metabelian group  $H_1^{H_2} \rtimes H_2$ , and we can complete the proof as above.  $\square$

**Theorem 4.18.** *For every recursive function  $f$ , there is a residually finite finitely presented solvable of class 3 group  $G$  with Dehn function greater than  $f$ . In addition, one can assume that the word problem in  $G$  is at least as hard as the membership problem in a given recursive set of natural numbers  $Z$  or as easy as polynomial time.*

*Proof.* The statement follows from Theorems 4.17 and 3.7.  $\square$

### 4.4. Residually finite finitely presented group with large depth function.

**Theorem 4.19.** *For every recursive function  $f$  there exists a finitely presented residually finite group  $G$  from  $\mathcal{A}_p^2 \cap \mathcal{ZN}_3 \mathcal{A}$  such that  $\rho_G(n) > f(n)$  for all  $n$ . In addition, we can assume that the word problem in  $G$  is as hard as the membership problem for any prescribed recursive set of natural numbers.*



*Proof.* Consider the Minsky machine  $M_n$  constructed in the proof of Theorem 3.9. Then as in the proof of Theorem 4.18, one can prove that  $G(M_n)$  is residually finite. The fact that  $\rho_G(n) > f(n)$  is proved the same way as in the proof of Theorem 3.9 (one only needs to replace the product by operation  $*$  everywhere in that proof).  $\square$

## 5. DISTORTION OF SUBGROUPS CLOSED IN THE PRO-FINITE TOPOLOGY

Let us generalize the Mikhailova construction [43].

Let  $G$  be a finitely generated group generated by a finite set  $X$ ,  $N \leq G$  a normal subgroup, generated as a normal subgroup by a finite set  $R = \{r_1, \dots, r_k\}$ , and  $\phi : G \rightarrow G/N$  the canonical epimorphism. We may assume that both sets  $X$  and  $R$  are symmetric, i.e.,  $X = X^{-1}$  and  $R = R^{-1}$ . The set

$$E(G, N) = \{(u, v) \in G \times G \mid \phi(u) = \phi(v)\}$$

is a subgroup of  $G \times G$ , called the *equalizer* of  $(\phi, \phi)$ .

In the following lemma we summarize the main components of Mikhailova's argument (though in a much more general situation). The proof is easy and we leave it to the reader.

**Lemma 5.1.** *In the notation above the following hold:*

- $E(G, N)$  is generated by a finite set

$$D = \{(r, 1) \mid r \in R\} \cup \{(x, x^{-1}) \mid x \in X\} \subset G \times G.$$

- For any  $w \in G$  if  $(w, 1) = (u_1, v_1)(u_2, v_2) \dots (u_n, v_n)$  for some  $(u_i, v_i) \in D$ , then  $u_1 \dots u_n$  is of the form  $w_0 r_1 w_1 r_2 w_2 \dots w_{m-1} r_m w_m$  for some  $w_i \in G$ ,  $r_i \in R$ ,  $m \leq n$ , satisfying  $w_0 w_1 \dots w_m = 1$  in  $G$ , hence,

$$w =_G \prod_{i=1}^m (w_0 \dots w_{i-1}) r_i (w_0 \dots w_{i-1})^{-1}.$$

*In particular, the distortion of  $E(G, N)$  in  $G \times G$  is at least as high as the Dehn function of  $G/N$  relative to  $N$ .*

Let  $\mathcal{P}$  be a class of finite groups closed under direct products and subgroups. Recall that the pro- $\mathcal{P}$  topology on a group  $G$  has as its base the set of all normal subgroups  $N$  with  $G/N \in \mathcal{P}$ .

**Lemma 5.2.** *Let  $\mathcal{P}$  be a class of finite groups closed under direct products and subgroups. In the notation above if the group  $G/N$  is residually  $\mathcal{P}$  then the subgroup  $E(G, N)$  is closed in the pro- $\mathcal{P}$  topology on  $G \times G$ .*

*Proof.* Suppose  $(u, v) \in G \times G$  but  $(u, v) \notin E(G, N)$ , so  $\phi(u) \neq \phi(v)$ . Since  $G/N$  is residually  $\mathcal{P}$  there is a homomorphism  $\eta : G/N \rightarrow K$  onto a finite group  $K \in \mathcal{P}$  such that  $\eta\phi(u) \neq \eta\phi(v)$  in  $K$ . Therefore the image of the pair  $(u, v)$  under  $\eta\phi$  is not in the image of the subgroup  $E(G, N)$  in  $K \times K$ . Hence the subgroup  $E(G, N)$  is closed in the pro- $\mathcal{P}$  topology on  $G \times G$ .  $\square$

The same argument gives the following

**Lemma 5.3.** *Under the assumptions of Lemma 5.2, the relative depth function  $\rho_{E(G, N)}$  is at least as large as the depth function of  $G/N$ , the time complexities of the “yes” and “no” parts of the membership problem for  $E(G, N)$  are as high as the time complexities of the “yes” and “no” parts of the word problem in  $G/N$ .*

Lemma 5.3 and Theorem 4.18 imply

*Remark 5.4.* The converse of Lemma 5.2 also holds, namely, if  $E(G, N)$  is closed in the pro- $\mathcal{P}$ -topology, then  $G/N$  is residually  $\mathcal{P}$ . We are not using this remark below so we leave it as an (easy) exercise.

**Theorem 5.5.** *For any recursive function  $f(n)$  there is a finitely generated subgroup  $H \leq F_2 \times F_2$  such that  $H$  is closed in the pro-finite topology on  $F_2 \times F_2$  and has distortion at least  $f(n)$ .*

*Proof.* Let  $G = \langle X \mid R \rangle$  be a finitely presented residually finite group with Dehn function at least  $f(n)$  from Theorem 4.18. If  $N$  is the normal closure of  $R$  in  $F(X)$  then the subgroup  $H = E(F(X), N) \leq F(X) \times F(X)$  satisfies all the requirements of the theorem.

Now one can embed the free group  $F(X)$  into  $F_2$  in such a way that the pro-finite topology induced on the image of  $F(X)$  from  $F_2$  is precisely the pro-finite topology on  $F(X)$ . Indeed, there is a finite index subgroup  $H$  of  $F_2$  of rank  $|X|$ , the induced topology on  $H$  is the pro-finite topology on  $H$ . It follows that the pro-finite topology on the subgroup  $F_{|X|}$  of  $F_2$  is precisely the topology induced by the pro-finite topology from  $F_2$ , as required.  $\square$

Applying the same argument to the free solvable groups  $S_3(X)$  of class 3 and generating set  $X$  one gets the following result.

**Theorem 5.6.** *For any recursive function  $f(n)$  there is a finite set  $X$  and a finitely generated subgroup  $H \leq S_3(X) \times S_3(X)$  such that  $E$  is closed in the pro-finite topology on  $S_3(X) \times S_3(X)$  and has distortion function, relative depth function, the time complexities of both “yes” and “no” parts of the membership problem and at least  $f(n)$ .*

## 6. UNIVERSAL THEORIES OF SETS OF FINITE SOLVABLE GROUPS

In this section we will prove the following result. For the class of all finite groups in was proved by Slobodskoi [60] (the idea of Slobodskoi’s proof came from Gurevich’s paper [23] where the same result was proved for semigroups).

**Theorem 6.1.** *The universal theories of the class of finite groups from  $\mathcal{A}_p^2 \mathcal{A} \cap \mathcal{ZN}_5 \mathcal{A}$  and the class of all periodic groups are recursively inseparable. In particular, the universal theory of any set of finite groups containing all finite solvable of class 3 groups is undecidable.*

*Proof.* It is well known [23] that there exists a Turing machine for which the set of input configurations accepted by the machine and the set of input configurations starting with which the machine never stops are recursively inseparable. Let  $M$  be a 2-glass Minsky machine with the same property.

Consider the 4-glass Minsky machine  $M_n$  described in the proof of Theorem 3.9. Let  $S'(M_n)$  be the semigroup given by the same defining relations as  $S(M_n)$  except the relation  $q_0 = 0$  is substituted by the relation  $q_i A_3 A_4 = 0$  for every  $i$ . It does not affect the proof of Theorem 4.3.

Let  $G'(M_n)$  be the group corresponding to  $S'(M_n)$  in the same way  $G(M_n)$  corresponds to  $S(M_n)$ . Then  $G'(M_n)$  belongs to  $\mathcal{A}_p^2 \mathcal{A} \cap \mathcal{ZN}_5 \mathcal{A}$  and simulates  $M_n$  as described in Theorem 4.3. Let  $R$  be the (finite) set of defining relations of  $G'(M_n)$ . Let  $X$  be the set of numbers  $\epsilon$  such that  $M_n$  accepts the configuration  $(\epsilon, 0, 1, 0)$ . Let  $X'$  be the set of numbers  $\epsilon$  such that  $M_n$  works infinitely long starting with the configuration  $(\epsilon, 0, 1, 0)$ . Then  $X$  and  $X'$

are recursively inseparable by the choice of  $M$  and  $M_n$ . For any configuration  $(\epsilon, 0, 1, 0)$  of  $M_n$  consider the corresponding element

$$w(\epsilon) = q_1 * a_1^{(\epsilon)} * A_1 * A_2 * a_3 * A_3 * A_4.$$

Suppose  $\epsilon \in X$ . Then there are only finite number of computations of  $\text{Sym}(M_n)$  starting at the configuration  $(1; \epsilon, 0, 1, 0)$ . Then as in the proof of Theorem 4.19, there exists a finite homomorphic image  $H$  of  $G'(M_n)$  where the image of  $w(\epsilon)$  is not equal to 1. Hence the universal formula  $\&R \rightarrow w(\epsilon) = 1$  does not hold in the finite group  $H$  from  $\mathcal{A}_p^2 \mathcal{A} \cap \mathcal{ZN}_5 \mathcal{A}$ .

Now suppose that  $\epsilon \notin X$ . Consider any periodic homomorphic image  $H$  of  $G'(M_n)$ . Let  $\bar{t}$  be the image of  $t \in G'(M_n)$  in  $H$ . Then there exists a number  $D$  such that for every element  $x \in \bar{T}$ ,

$$(17) \quad x * \bar{a}_3^{(D)} = x * \bar{a}_3^{(2D)}.$$

Since  $M_n$  works infinitely long starting at the configuration  $(\epsilon, 0, 1, 0)$ , by Theorem 4.3 the following equality is true for some  $i, k_1, k_2$ :

$$w(\epsilon) = \bar{x}_{1, q_i A_0} * \bar{a}_1^{(k_1)} * \bar{a}_2^{(k_2)} * \bar{a}_3^{(D)} * \bar{A}_1 * \bar{A}_2 * \bar{A}_3 * \bar{A}_4.$$

Then by (17) and Theorem 4.3

$$\begin{aligned} \bar{w}(\epsilon) &= \bar{x}_{1, q_i A_0} * \bar{a}_1^{(k_1)} * \bar{a}_2^{(k_2)} * \bar{a}_3^{(2D)} * \bar{a}_4^{(D)} * \bar{A}_1 * \bar{A}_2 * \bar{A}_3 * \bar{A}_4 \\ &= \bar{x}_{1, q_i A_0} * \bar{a}_1^{(k_1)} * \bar{a}_2^{(k_2)} * \bar{a}_3^{(D)} * \bar{a}_4^{(D)} * \bar{A}_1 * \bar{A}_2 * \bar{A}_3 * \bar{A}_4 \\ &= \bar{x}_{1, q_i A_0} * \bar{a}_1^{(k_1)} * \bar{a}_2^{(k_2)} * \bar{A}_1 * \bar{A}_2 * \bar{A}_3 * \bar{A}_4 = 1. \end{aligned}$$

since  $q_i A_3 A_4 = 0$  in  $S'(M_n)$ . Hence the universal formula  $\&R \rightarrow w(\epsilon) = 1$  holds in  $H$ .

Thus the set of universal formulas  $\&R \rightarrow w(\epsilon) = 1$  that do not hold in some finite group from  $\mathcal{A}_p^2 \mathcal{A} \cap \mathcal{ZN}_5 \mathcal{A}$  and the set of such formulas which hold in every periodic group are recursively inseparable.  $\square$

*Remark 6.2.* Note that the universal theory of finite metabelian groups is decidable [32]. The same is true for the set of finite groups (and any other algebraic structures of finite type) of any finitely based variety where every finitely generated group is residually finite [32]. On the other hand, the universal theory of all finite nilpotent groups is undecidable [30]. The description of all (finitely based) varieties of groups where the universal theory of finite groups is decidable is currently out of reach. From Zelmanov's solution of the restricted Burnside problem [65, 66], it immediately follows that the universal theory of finite groups in every finitely based periodic variety of groups is decidable. That result and simulations of Minsky machines in semigroups (as in Section 3) were used by the third author [56] to obtain a complete description of all finitely based varieties of semigroups where finite semigroups have decidable universal theory. For more information on that problem, see [32].

## REFERENCES

- [1] I. Agol, The virtual Haken conjecture (with an appendix by I. Agol, D. Groves and J. Manning), arXiv:1204.2810, 2012.
- [2] G. Baumslag, A non-cyclic one-relator group all of whose finite factor groups are cyclic, J. Australian Math. Soc., 10, pp. 497–498, (1969).
- [3] G. Baumslag, Subgroups of finitely presented metabelian groups. J. Austr. Math. Soc. (Series A), 16(1):98–110, 1973.

- [4] G. Baumslag, C.F. Miller III, H. Short, Isoperimetric inequalities and the homology of groups, *Invent. Math.*, 113 (3) (1993), pp. 531–560.
- [5] G. Baumslag and J. E. Roseblade, Subgroups of direct products of free groups, *J. London Math. Soc.* (2), 30 (1984), 44–52.
- [6] Bestvina, M.; Feighn, M. A combination theorem for negatively curved groups, *J. Differential Geom.*, 35 (1992), no. 1, 85–101.
- [7] Alexander Borisov, Mark Sapir, Polynomial maps over finite fields and residual finiteness of mapping tori of group endomorphisms. *Invent. Math.* 160 (2005), 2, 341–356.
- [8] Alexander Borisov, Mark Sapir, Polynomial maps over  $p$ -adics and residual properties of mapping tori of group endomorphisms. *Int. Math. Res. Not. IMRN* 2009, 16, 3002–3015.
- [9] K. Bou-Rabee, Quantifying residual finiteness, *Journal of Algebra*, 323 (2010) 729–737.
- [10] Noel Brady, Will Dison, Tim Riley, Hyperbolic hydra, [arXiv:1105.1535](https://arxiv.org/abs/1105.1535).
- [11] M. Bridson, A. Haefliger. *Metric spaces of non-positive curvature*, Springer, 1999.
- [12] Daniel E. Cohen, *Combinatorial group theory: a topological approach*. London Mathematical Society Student Texts, 14. Cambridge University Press, Cambridge, 1989.
- [13] M. D. Davis, A note on universal Turing machines. *Automata studies*, pp. 167–175. *Annals of mathematics studies*, no. 34. Princeton University Press, Princeton, N. J., 1956.
- [14] Will Dison, Tim Riley, Hydra groups, [arXiv:1002.1945](https://arxiv.org/abs/1002.1945).
- [15] Verena Huber Dyson, A family of groups with nice word problems. Collection of articles dedicated to the memory of Hanna Neumann, VIII. *J. Austral. Math. Soc.* 17 (1974), 414–425.
- [16] Benson Farb, The extrinsic geometry of subgroups and the generalized word problem. *Proc. London Math. Soc.* (3) 68 (1994), no. 3, 577–593.
- [17] M. R. Garey, D. S. Johnson, *Computers and intractability. A guide to the theory of NP-completeness*. A Series of Books in the Mathematical Sciences. W. H. Freeman and Co., San Francisco, Calif., 1979.
- [18] S. M. Gersten, Dehn functions and  $l_1$ -norms of finite presentations. *Algorithms and Classification in Combinatorial Group Theory*, pp. 195–225. Springer, Berlin, 1992.
- [19] S. M. Gersten, Isoperimetric and isodiametric functions of finite presentations. *Geometric group theory*, Vol. 1 (Sussex, 1991), 79–96, *London Math. Soc. Lecture Note Ser.*, 181, Cambridge Univ. Press, Cambridge, 1993.
- [20] Steve M. Gersten and Tim R. Riley, Some duality conjectures for finite graphs and their group theoretic consequences, *Proc. Edin. Math. Soc.*, 48(2), 2005, 389–421.
- [21] R. Grigorchuk, Groups with intermediate growth functions and their applications, Doctor’s Thesis (Russian), Moscow Steklov Mathematical Institute, 1985.
- [22] E. A. Golubov, Finite separability in semigroups. *Dokl. Akad. Nauk SSSR* 189 (1969), 20–22.
- [23] Ju. Sh. Gurevich, The problem of equality of words for certain classes of semigroups. *Algebra i Logika Sem.* 5 1966 no. 5, 25–35.
- [24] G. Higman, Subgroups of finitely presented groups. *Proc. Roy. Soc. Ser. A*, 262 (1961), 455–475.
- [25] Tim Hsu, Daniel Wise, Cubulating graphs of free groups with cyclic edge groups. *Amer. J. Math.* 132 (2010), no. 5, 1153–1188.
- [26] M. Kassabov, F. Matucci, Bounding the residual finiteness of free groups, preprint, [arXiv](https://arxiv.org/abs/1002.1945).
- [27] M. Kassabov, N. Nikolov, Generation of polycyclic groups. *J. Group Theory* 12 (2009), no. 4, 567–577.
- [28] O. G. Kharlampovich, Finitely presented solvable group with unsolvable word problem, *Soviet Math. Izvestia*, 45, 4, 1981, 852–873.
- [29] O. G. Kharlampovich, The word problem for groups and Lie algebras, Doctor’s Thesis (Russian), Moscow Steklov Mathematical Institute, 1990.
- [30] O. G. Kharlampovich, The universal theory of the class of finite nilpotent groups is undecidable. *Mat. Zametki* 33 (1983), no. 4, 499–516.
- [31] O.G. Kharlampovich, M.V. Sapir, A non-residually finite, relatively finitely presented group in the variety  $\mathfrak{N}_2\mathfrak{L}$ . *Combinatorial and geometric group theory* (Edinburgh, 1993), 184–189, *London Math. Soc. Lecture Note Ser.*, 204, Cambridge Univ. Press, Cambridge, 1995.
- [32] O. Kharlampovich, M. Sapir, Algorithmic problem in varieties, *International Journal of Algebra and Computation*, vol 5, no 4,5, 379–602, 1995.
- [33] B. Khoussainov, A. Myasnikov, On finitely presented expansions of semigroups, groups, and algebras. To appear in *Trans. of Amer. Math. Soc.*
- [34] Kourouvskaia tetrad’ (Unsolved problems in Group theory), 5-th edition. Novosibirsk, 1976.

- [35] R. J. Lipton and Y. Zalcstein, Word problems solvable in logspace, *J. Assoc. Comput. Mach.* 24 (1977), 522–526.
- [36] A. I. Malcev, *Algorithms and recursive functions*, Moscow, Nauka, 1965.
- [37] A. I. Malcev, On Homomorphisms onto finite groups (Russian). *Uchen. Zap. Ivanovskogo Gos. Ped. Inst.* 18 (1958), 49–60. English translation in: *Amer. Math. Soc. Transl. Ser. 2*, 119 (1983) 67–79.
- [38] Ralph McKenzie, Richard J. Thompson, An elementary construction of unsolvable word problems in group theory. Word problems: decision problems and the Burnside problem in group theory (Conf., Univ. California, Irvine, Calif. 1969; dedicated to Hanna Neumann), *Studies in Logic and the Foundations of Math.*, 71, pp. 457–478. North-Holland, Amsterdam, 1973.
- [39] S. Meskin, A Finitely Generated Residually Finite Group with an Unsolvable Word Problem, *Proceedings of the American Mathematical Society*, 43, 1 (1974), 8–10.
- [40] K. Madlener, F. Otto, Pseudonatural algorithms for the word problem for finitely presented monoids and groups, *J. Symbolic Comput.*, 1 (1985), no. 4, 383–418.
- [41] J. McKinsey, The decision problem for some classes of sentences without quantifiers, *J. Symbolic Logic* 8 (1973), 61–76.
- [42] A. Miasnikov, A. Ushakov, and D. Won, The Word Problem in Baumslag Group is polynomial time decidable, *Journal of Algebra* 345 (2011), pp. 324–342.
- [43] K. A. Mikhailova, The occurrence problem for direct products of groups, *Dokl. Akad. Nauk SSSR* 119 (1958), pp. 1103–1105.
- [44] A. Miasnikov, V. Roman'kov, A. Ushakov, and A. Vershik, The word and geodesic problems in free solvable groups, *Trans. Amer. Math. Soc.* 362 (2010), no. 9, 4655–4682.
- [45] H. Neumann, *Varieties of groups*. Springer-Verlag, Berlin, Heidelberg, 1967.
- [46] N. Nikolov, D. Segal, Finite index subgroups in pro-finite groups. *C. R. Math. Acad. Sci. Paris* 337 (2003), 5, 303–308.
- [47] Yann Ollivier, Daniel T. Wise, Cubulating random groups at density less than  $1/6$ . *Trans. Amer. Math. Soc.* 363 (2011), no. 9, 4701–4733.
- [48] A. Yu. Olshanskii, Almost every group is hyperbolic. *International Journal of Algebra and Computation* 2 (1992), no. 1, 1–17.
- [49] A. Olshanskii and M. Sapir, Length and area functions on groups and quasi-isometric Higman embeddings, *International Journal of Algebra and Computation* 11 (2001), 137–170.
- [50] W. Parry, Growth Series of Some Wreath Products, *Trans. Am. Math. Soc.*, 331, 2, 1992, 751–759.
- [51] A.N. Platonov, An isoperametric function of the Baumslag-Gersten group. *Vestnik Moskov. Univ. Ser. I Mat. Mekh.* 2004, no. 3, 12–17, translation in *Moscow Univ. Math. Bull.* 59 (2004), no. 3, 1217 (2005).
- [52] V. Remeslennikov, Studies on infinite solvable and finitely approximable groups. *Mat. Zametki* 17 (1975), no. 5, 819–824.
- [53] E. Rips, Subgroups of small cancellation groups, *Bull London Math Soc.*, 14 (1982), 45–47.
- [54] Joseph J. Rotman, *An introduction to the theory of groups*. Fourth edition. Graduate Texts in Mathematics, 148. Springer-Verlag, New York, 1995. xvi+513 pp.
- [55] Mark Sapir, Algorithmic problems in varieties of semigroups. *Algebra i Logika* 27 (1988), 4, 440–463.
- [56] Mark Sapir, Weak word problem for finite semigroups. *Monoids and semigroups with applications* (Berkeley, CA, 1989), 206–219, World Sci. Publ., River Edge, NJ, 1991.
- [57] Mark Sapir, Asymptotic invariants, complexity of groups and related problems. *Bull. Math. Sci.* 1 (2011), no. 2, 277–364.
- [58] Mark Sapir, Jean-Camille Birget, Eliyahu Rips, Isoperimetric and isodiametric functions of groups. *Ann. of Math.* (2) 156 (2002), 2, 345–466.
- [59] H. U. Simon, Word problems for groups and contextfree recognition, in *Fundamentals of computation theory* (Proc. Conf. Algebraic, Arith. and Categorical Methods in Comput. Theory, Berlin/Wendisch-Rietz, 1979), Akademie-Verlag, Berlin, (1979), 417–422.
- [60] A.M. Slobodskoi, Undecidability of the universal theory of finite groups, *Algebra and Logic*, 1981, 20, 2, 207–230.
- [61] St. Waack, On the parallel complexity of linear groups. *RAIRO Inform. Theor. Appl.* 25 (1991), 323–354.
- [62] B. A. F. Wehrfritz, On finitely generated soluble linear groups. *Math. Z.* 170 (1980), no. 2, 155–167.
- [63] D. T. Wise, The structure of groups with a quasiconvex hierarchy, preprint, 2011.
- [64] D. T. Wise, A residually finite version of Rips's construction. *Bull. London Math. Soc.* 35 (2003), no. 1, 23–29.

- [65] E. I. Zel'manov, The solution of the restricted Burnside problem for groups of odd exponent. *Izv. Akad. Nauk. SSSR. Ser. Mat.*, 54(1):42–59, 1990. transl. in *Math. USSR-Izv.* 36 (1991), no.1, 41-60.
- [66] E. I. Zel'manov, The solution of the restricted Burnside problem for 2-groups. *Mat. Sb.*, 182(4):568–592, 1991.

OLGA KHARLAMPOVICH, DEPARTMENT OF MATHEMATICS AND STATISTICS, HUNTER COLLEGE, CITY UNIVERSITY OF NEW YORK, NEW YORK, NY, 10065, U.S.A.

*E-mail address:* okharlampovich@gmail.com

ALEXEI MYASNIKOV, STEVENS INSTITUTE OF TECHNOLOGY, HOBOKEN, NJ, 07030 U.S.A.

*E-mail address:* amiasnik@stevens.edu

MARK SAPIR, DEPARTMENT OF MATHEMATICS, VANDERBILT UNIVERSITY, NASHVILLE, TN 37240, U.S.A.

*E-mail address:* m.sapir@vanderbilt.edu